Determinants of seed size and yield in Arabidopsis thaliana

Hughes, Rhiannon

Award date:
2009

Awarding institution:
University of Bath

Determinants of seed size and yield in *Arabidopsis thaliana*

Copy 1 of 3

Rhiannon Hughes

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

July 2009

Signed:

Rhiannon Hughes

For my Grandad Les

ACKNOWLEDGEMENTS

I would like to say a huge thank you to my supervisor Rod Scott for his never-ending patience, encouragement and enthusiasm during my PhD. A special thank you is also due to Melissa Spielman for her unwavering support in the lab and significant contribution to my first publication. I am also very grateful to Sushma Tiwari for sharing her wealth of knowledge and experience of molecular biology, and giving me the confidence to keep going when things didn't go quite as planned. Thank you to all past and present members of lab 3S 1.13 for their support and friendship: Pear, Lihua, Marie, Bolbol, Ahlam, Valentyna, Lindsay, Lucy, Baoxiu, Sapu, Richard, Alec and Holly.

My grateful thanks are given to the following: Marie Schruff for her work on the *arf2* project before I started and introducing me to Emma and Betsy; Ian Graham and Tony Larson for analysing the lipid content of *arf2* seeds; Angie Vale and Julia Watling for caring for my plants; Mike Mogie for help with statistics; Lindsay Dytham for her fun company, girly chats and great nights out dancing in Canada 'on conference'; Kimbo for getting me a place on the Equestrian Team, letting Henna throw me off and everything else horse-related; Rosie for getting married so we all had something to get excited about; the Society of Experimental Biology for funding my attendance at conferences; and the BBSRC for funding this research.

Appreciation must be extended to those who provided temporary accommodation: Emma and her family (Bath); my Mum and Dad (Hereford); Rosie and Matt (Bradford-on-Avon); Gill and Tim (Rotherfield); Gavin and Laura (Bristol); and Kimbo and her parents (Bath).

Lastly and most importantly I want to thank Mike, not only for his technical advice and the many hours spent labelling tubes and typing in data, but also for taking, out, my, superfluous, commas!

ABSTRACT

It is becoming increasingly important to improve the yield of seed crops to feed an expanding population and, more recently, to cope with additional strains on food-oriented agriculture posed by biofuel production and global climate change. One strategy to increase yield is to increase seed size. However in nature, due to the limited resources of the mother plant, an enlarged seed size is often associated with a concomitant reduction in seed number. Using the model plant *Arabidopsis thaliana*, a biotechnological approach to increase seed size through the modification of the triploid endosperm or the ovule integuments was shown to be a viable strategy to improve seed yield. Targeted over-expression of the *ANT* transcription factor specifically within the endosperm significantly increased seed size without negative effects on fertility. Overcoming compromised fertility in *arf2* mutants established that *ARF2* and the integuments are important factors in determining seed size and yield. To ensure agronomic relevance, both the gross yield and the harvest index (HI, ratio of seed yield to biological yield) were used to assess the impact of traits, such as increased integument size, introduced into *Arabidopsis*.

To uncover novel regulators of seed size and further develop the current understanding of seed development, second-site mutations were induced in the *auxin response factor 2* (*arf2*) mutant, which produces large seeds due to extra cell division in the ovule integuments. The *ARABIDOPSIS CRINKLY 4* (*ACR4*) gene was identified as a suppressor of the *arf2* seed phenotype. In *arf2-8 acr4* double mutants, an additive effect on cell proliferation was observed, indicating that *ACR4* affects the seed coat through a developmental pathway independent of *ARF2*.

Natural variation present in *Arabidopsis* was used to study seed yield and its components. Considerable variation in seed size, yield and HI was revealed. Significantly, high seed weight was not associated with high yield or high HI. In contrast, high seed number and reduced plant stature were revealed as important components of high yield and yield efficiency. Additionally, *ERECTA* was identified as a potential 'Green Revolution' gene. Seed size in

Arabidopsis **was shown to be an extremely plastic trait in response to alterations in the post-flowering source-sink ratio, increasing by 35.0 – 66.2% in the ecotypes tested. Furthermore, a difference in seed weight between restricted (high source-sink ratio) and unrestricted (low source-sink ratio) pollinations was first observed remarkably early in seed development. However, reducing seed number by restricting pollination did not substantially alter integument or endosperm development in order to facilitate increased nutrient uptake.**

ABBREVIATIONS & ACRONYMS

ACR4	ARABIDOPSIS CRINKLY 4
AGL28	AGAMOUS-LIKE 28
ANOVA	analysis of variance
ANT	AINTEGUMENTA
AP1	APETALA 1
AP2	APETALA 2
ARF2	AUXIN RESPONSE FACTOR 2
Bla	Blanes/Gerona
cds	coding sequence
Col	Columbia
Cvi	Cape Verde Islands
DAP	days after pollination
DNA	deoxyribonucleic acid
EMS	ethylmethane sulphonate
ER	ERECTA
FA	fatty acid
FAME	fatty acid methyl ester
GUS	ß-glucuronidase
HI	harvest index
ii1	the innermost layer of the inner integument (endothelium)
ii1'	a layer of the inner integument that spans only part of the embryo sac
ii2	the outer layer of the inner integument
INO	INNER NO OUTER
Kas	Kashmir
L*er*	Landsberg *erecta* (mutant)
L*ER*	Landsberg *ERECTA* (wild type)
MNT	MEGAINTEGUMENTA
NASC	Nottingham Arabidopsis Stock Centre
Ob	Oberursel/Hasen
ocs	octopine synthase terminator
oi1	the inner layer of the outer integument
oi2	the outer layer of the outer integument (epidermis)
pCZE	At5g07210 chalazal endosperm-specific promoter
PHE1	PHERES1
pPER	At5g46950 endosperm-specific promoter
RNA	ribonucleic acid
RNAi	RNA interference
Rsch	Rschew/Starize
s.e.m.	standard error of the mean
SFD	seed fill duration
SGR	seed growth rate
SHP2	SHATTERPROOF 2

CONTENTS

1. INTRODUCTION

1.1 Why study seed size and yield?

1.1.1 Seeds provide the bulk of our diet

The world's food supply is entirely dependent on the ability of photosynthetic plants to convert sunlight into useable chemical energy. Humans cultivate plants for their fruits, roots, tubers and vegetative organs; however undoubtedly the bulk of our diet is derived from plants harvested for their mature seed. Seeds are extremely rich in carbohydrate, protein and oil, reserves the next plant generation requires to become established, but also a vital source of nutrition for both humans and animals. Although a huge variety of seed crops is grown, global production of cereals, mainly wheat, rice and maize, is far greater than for any other crops. The UK alone produces about 15.5 million tonnes of wheat per year, and cereals are grown on approximately 70% of all arable land (DEFRA, 2008). However, the UK represented less than 1% of the total worldwide cereal production in 2007 (FAO, 2007). Seeds therefore have great economic value and are fundamental for world food security.

1.1.2 Increasing demand and past solutions

World population increased dramatically from 3 billion to 6 billion between 1959 and 1999, and is predicted to reach more than 9 billion by 2050 (IDB, 2008). In the past, seed crop yields have increased alongside population growth; however there is concern that this parallel cannot continue. Despite a reduction in world hunger over the past 30 years, in 2002 there were more than 800 million malnourished people in the developing world, including more than a quarter of children under age 5 (FAO, 2002). Crop production must not only be increased to feed an expanding population but also to allow for changing global diets. The consumption of animal-based foods, particularly in developing countries where more than 85% of the population resides, is rising dramatically (Shapouri and Rosen, 2007). This demands a higher energy input per person compared to plant-based diets (Eshel and Martin, 2005). Approximately 40% of the world's land surface is already used for agricultural purposes (Foley *et al*., 2005). Therefore, sequestering more land for crop production is not a viable solution, especially given the recent competition from growing crops for biofuels (Hill, 2007).

Enhancing the yield of seed crops may be achieved by increasing resistance to pests and diseases, improving tolerance to adverse environmental conditions, improving farming practices and increasing genetic yield potential. Dramatic yield improvements were

achieved during the 'Green Revolution' in the 1960s with the introduction of dwarf cereal varieties. Conventional breeding led to the development of short-stemmed wheat and rice plants that were lodging resistant when grown using high levels of fertilizer, therefore yield losses were reduced. However, from the crop physiologist's perspective, the more remarkable improvement was the rise in the harvest index (HI). The HI has been defined as the ratio of (wheat) grain yield to biological yield or biomass (Donald, 1962; Donald and Hamblin, 1976). Higher HI ratios indicated that the proportion of plant biomass allocated to the harvestable plant parts, in this case the seeds, had increased. In the 1920s the HI of wheat was roughly 0.35, whereas modern varieties achieve from 0.50 to 0.55, producing as much grain as straw (Evans, 1993). The superior cereal varieties were central to the success of the Green Revolution; however increased agronomic support through irrigation and the application of fertilizer, herbicides and pesticides also contributed greatly. These inputs are so critical that it has been suggested that without them it would not be possible to support the present population, as high yields can only be sustained with high inputs (Evans, 1993).

Seed yield has often been considered too complex a trait to be influenced by the modification of a single gene. This implies that the genetic engineering of new, high-yielding varieties will be difficult. Genetically modified (GM) crops were first developed in the 1980s and released in the mid 1990s, and primarily were developed to confer insect resistance or herbicide tolerance (Huang *et al*., 2002). By 2005, transgenic crops were grown on an estimated 90 million hectares by 8.5 million farmers in 21 countries (James, 2005). Recent advances in biotechnology and improved understanding of yield at the molecular level suggest that there is now the potential to target yield directly. Increased total seed weight per plant has been achieved by modifying single genes involved in starch biosynthesis (Smidansky *et al*., 2002, 2003), assimilate partitioning (Heyer *et al*., 2004) and plant architecture (Ait-Ali *et al*., 2003). A further approach to enhance seed yield is to manipulate the genes involved in seed development with the aim of increasing seed size.

1.1.3 A potential solution: increasing seed size to enhance yield

The seed yield of crop plants is estimated as the seed weight per unit area, i.e. tonnes per hectare, therefore seed size is an important component of yield. Since agriculture began, seeds have been subjected to selection and breeding for increased size and other qualities, and consequently crops today produce much larger grains than their wild relatives (Sasahara, 1984; Evans, 1993). Large seeds are positively associated with germination, seedling establishment and survival (Krannitz *et al*., 1991; Westoby *et al*., 1996; Koelewijn and Van Damme, 2005), and ease of processing (Tanska *et al*., 2008). Therefore, the

manipulation of seed size by genetic engineering is an attractive strategy for many reasons, not solely to increase seed yield. Seed size among different plant species varies over a range of ten orders of magnitude, from the extremely small, as borne on the orchid *Goodyera repens* (2 μg), to the enormous seeds of the double coconut palm *Lodoicea maldivica* (27 kg) (Harper *et al*., 1970). Therefore, huge genetic potential for seed size manipulation exists. However, seed size within a particular species can be remarkably stable; so consistent that the seed of the Carob tree *Ceratonia siliqua* was originally used to counterweight gold (Harper *et al*., 1970). Seed size stability therefore presents a challenge that biotechnology must overcome in order to raise maximum seed yields with this approach. Achieving this goal requires a comprehensive understanding of the molecular and genetic regulation of seed development and seed size.

1.2 Advantages of the model plant *Arabidopsis thaliana*

The model plant *Arabidopsis thaliana* is a member of the mustard (Brassicaceae) family. Although considered a weed, it has proved invaluable as a research tool for unravelling the cellular and molecular biology of flowering plants (Somerville and Koorneef, 2002). *Arabidopsis* has many attributes which make it a good model plant, the most notable being its relatively small genome size. The *Arabidopsis* genome is estimated to be approximately 130 Mb in size and stores the genetic codes for roughly 25,000 genes. In comparison, many crop species have extremely large genomes due to polyploidization events and the accumulation of non-coding sequences, making the isolation, cloning and sequencing of mutant loci difficult. The genome size of oilseed rape, *Brassica napus*, is roughly 1,200 Mb and wheat, *Triticum aestivum*, is even larger, consisting of 16,000 Mb (Arumuganathan and Earle, 1991). The small genome of *Arabidopsis* facilitated the elucidation of the entire DNA sequence (The Arabidopsis Genome Initiative, 2000), which is now available from the TAIR website (www.arabidopsis.org). This website provides a wealth of information to the *Arabidopsis* user with details concerning genes, markers, polymorphisms, maps, clones, gene families and proteins.

Arabidopsis users also benefit from seed stock centres, such as the Nottingham Arabidopsis Stock Centre (NASC), that hold large collections of naturally-occurring accessions (ecotypes) and mutant lines. *Arabidopsis* ecotypes, collected from locations across the northern hemisphere, show wide variation in many physiological traits and thus provide an extremely useful resource for identifying the molecular basis of complex traits. An extensive collection of T-DNA insertion mutants is also available (Alonso *et al*., 2003) enabling the role of many genes with unknown function to be identified.

Other advantages of *Arabidopsis* include the ease of transformation using *Agrobacterium tumefaciens* and efficient mutagenesis with the chemical mutagen EMS. *Arabidopsis* is also a small plant that is easy to grow, completing its lifecycle in 6-8 weeks. It is self-fertile and highly prolific, producing thousands of seeds from an individual plant. Importantly for this research, the seed development of *Arabidopsis* is well characterized. *Arabidopsis* is also closely related to the commercially important *B. napus*; the coding regions of their genes show approximately 87% nucleotide identity (Cavell *et al*., 1998). Consequently, *Arabidopsis* is an extremely useful model for oilseed crops.

1.3 Seed development in *Arabidopsis*

Seed development encompasses such a large area of research that it would be impossible to cover in its entirety here. Many events occur long before fertilisation, such as the development of the ovule, but these cannot be overlooked as ovules are the precursors of seeds and the site of female gametogenesis and fertilisation. After fertilisation, the seed structures are formed through rapid cell division, followed by cellular expansion and differentiation. Embryogenesis, endosperm and seed coat development will be described in this chapter. As this thesis is concerned with seed size, i.e. the extent of seed fill, an explanation of phloem-unloading and seed storage compound accumulation will also be included.

1.3.1 Ovule development

A detailed morphological analysis of ovule development has been carried out by Schneitz *et al*. (1995). Mature ovules consist of three main structures: the embryo sac, the integuments and a stalk-like structure, known as the funiculus, which connects the ovule to the placental region of the gynoecium (Figure 1.1). The funiculus is the only ovular component which does not form part of the seed. While the integuments and funiculus are both diploid and therefore sporophytic, the embryo sac is haploid, and represents the female gametophytic phase of the plants lifecycle.

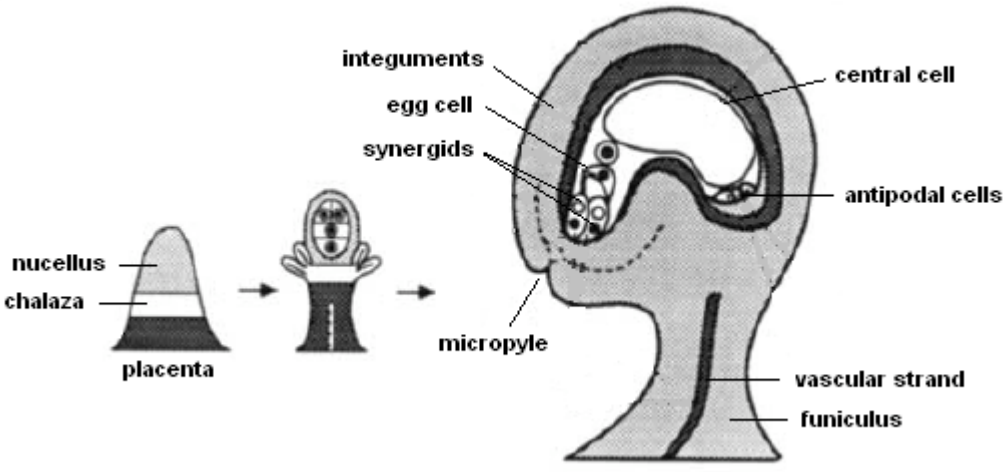

Figure 1.1: Schematic of ovule development in *Arabidopsis thaliana*.
Primordia arise on the placenta and differentiate into three distinct zones: the nucellus, chalazal and the proximal region which gives rise to the funiculus. Within the nucellus, female gametogenesis occurs culminating in a seven-celled, eight-nucleate embryo sac. Integuments initiate from the chalaza, and eventually surround the entire embryo sac except for a small opening, termed the micropyle. Adapted from Schneitz *et al*. (1995).

Ovule development takes place inside the gynoecium, where ovule primordia initiate from the placenta. The tip of a primordium differentiates into the nucellus, inside which the megasporocyte undergoes meiosis to produce four haploid cells. Only the proximal-most megaspore survives and develops into the embryo sac. The central domain of a primordium, known as the chalaza, is where the integuments initiate (Robinson-Beers *et al*., 1992). The asymmetric growth of the integuments gives the ovule its characteristic curvature. The integuments enclose the nucellus, leaving a small opening, the micropyle, through which the pollen tube may enter and deliver its cargo of sperm. The proximal region of a primordium develops into the funiculus. In the final stage of ovule development, the vascular strand appears within the funiculus, ready to provide water and nutrients for seed development following fertilisation.

Arabidopsis ovules are of *Polygonum*-type (Reiser and Fischer, 1993), each containing a mature gametophyte that has eight nuclei within seven cells. The egg apparatus is situated at the micropylar end of the embryo sac and consists of two synergids and the egg cell proper. The majority of the embryo sac is occupied by the central cell which carries a single dihaploid nucleus formed by the fusion of the two polar nuclei. At the chalazal end of the embryo sac, three antipodal cells are found, however these usually degenerate prior to or during fertilisation.

1.3.2 Double fertilisation

Uniquely, reproduction in angiosperms, including *Arabidopsis*, requires a double fertilisation event. Pollen grains fortunate enough to land on a receptive stigma will germinate to produce a pollen tube that delivers two sperm cells to a single ovule. Upon fertilisation, the ovule is re-defined as a developing seed and the gynoecium as a silique (or seed pod). One sperm fertilises the egg forming the diploid embryo, and the other fuses with the two polar nuclei to form the triploid endosperm (Lopes and Larkins, 1993). These two fertilisation products are surrounded by the third seed component, the testa or seed coat which is derived from the ovule integuments and in contrast, is entirely of maternal origin.

1.3.3 Embryogenesis

Inside the developing seed, the zygote undergoes a complex series of morphological and cellular changes in a process known as embryogenesis (West and Harada, 1993). The first division of the zygote is asymmetrical, generating a small apical cell and a large basal cell. Subsequently, the apical cell undergoes two cell divisions giving rise to the four-celled embryo proper. Continued cell division establishes distinctive embryonic growth stages

which can be used as a guide during seed development: octant, globular, heart (Figure 1.2), torpedo, and walking stick stages. At heart stage, two cotyledons emerge and many different tissue types can be distinguished, such as the procambium (precursor of the vascular tissue), the ground meristem and the shoot and root apices. Finally, when the embryo fills the entire seed cavity it is considered to be mature.

While the apical cell forms the embryo proper, the basal cell undergoes a series of transverse divisions, resulting in the formation of the hypophysis and the suspensor (Figure 1.2). The hypophysis, the uppermost derivative of the basal cell, will develop into the root tissues. The suspensor is a transient structure which may have a role in pushing the embryo into the nutrient-rich endosperm or may serve as a conduit for the transfer of nutrients and growth factors from the mother plant to the embryo (Yeung and Meinke, 1993).

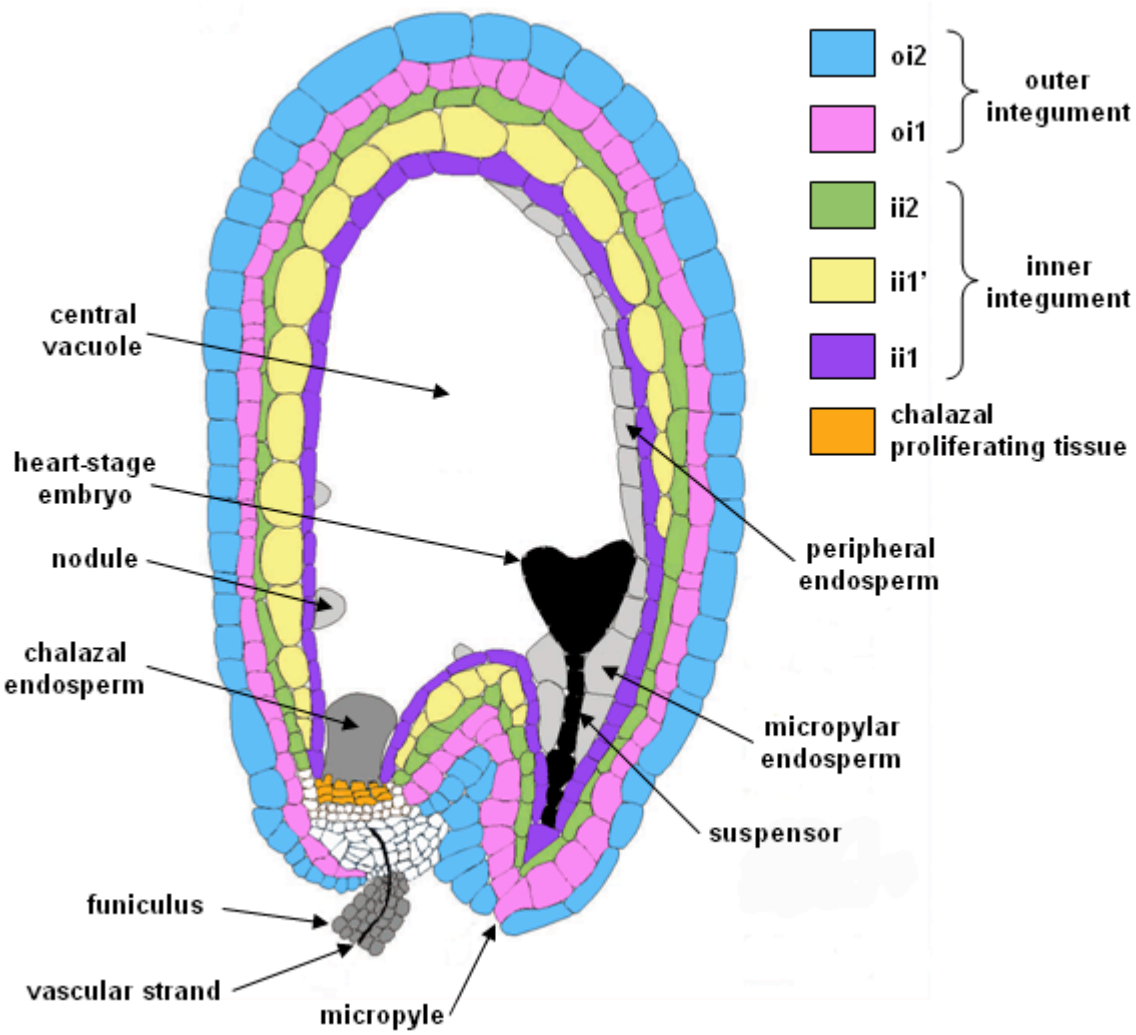

Figure 1.2: Diagram of seed structures in *Arabidopsis thaliana*.

Adapted from Debeaujon *et al*. (2003). Seed coat layers labelled according to Beeckman *et al*. (2000).

1.3.4 Endosperm development

Successful embryogenesis requires the development of the second fertilisation product, the triploid endosperm (Lopes and Larkins, 1993). The female gametophyte of non-flowering plants is much larger than that of angiosperms and provides plenty of nutrients for the developing embryo (Westoby and Rice, 1982). In contrast, angiosperms have little or no reserve food before fertilisation and therefore the principal function of the endosperm is typically regarded to be the synthesis and accumulation of storage products for the growing embryo. In *Arabidopsis*, and the majority of flowering plants endosperm development is nuclear, i.e. the primary endosperm nucleus undergoes several rounds of division without cell wall production, resulting in a large number of free nuclei within a syncytium (Lopes and Larkins, 1993). This developmental program enables the endosperm to grow more rapidly than the embryo which it will support. The free nuclei assemble at the periphery of the central cell and cytokinesis begins here and progresses towards the centre until the entire endosperm is cellular (Olsen, 2004). The endosperm can be classified into three main functional domains: the peripheral endosperm, the chalazal endosperm and the micropylar endosperm (Figure 1.2). The chalazal endosperm sits at the maternal-filial interface and is therefore thought to have a more specialised role in nutrient uptake. The micropylar endosperm may provide extra protection for the embryo which it surrounds or may be involved in communication between endosperm and embryo.

1.3.5 Seed coat development

Fertilisation triggers the differentiation of the ovule integuments into the seed coat (Beeckman *et al*., 2000). The seed coat has been assigned many important roles including assimilate transfer to the growing embryo (Stadler *et al*., 2005), protection of the embryo from mechanical damage or pathogen attack (Mohamed-Yasseen *et al*., 1994), maintenance of dormancy until favourable conditions for seedling growth exist and water uptake during germination (Debeaujon *et al*., 2000). The *Arabidopsis* seed coat consists of 5 layers: ii1, the inner layer of the inner integument; ii1′, a layer of the inner integument which spans only part of the embryo sac; ii2, the outer layer of the inner integument; oi1, the inner layer of the outer integument; and oi2, the outer layer of the outer integument (Figure 1.2, Beeckman *et al*., 2000). Following fertilization, all layers grow rapidly through both cell division and expansion. Cells of the ii1 layer, also referred to as the endothelium or pigment layer, synthesize flavonoid compounds called condensed tannins which give the seed coat its brown appearance. This layer may also supply nutrients to the embryo sac. Layers ii1’ and ii2 do not appear to differentiate further and eventually are crushed by the growing embryo. Both layers of the outer seed coat, the oi1 and oi2,

accumulate starch-containing amyloplasts before differentiating further (Windsor *et al*., 2000). The oi1 then produces a thickened cell wall and cells of the oi2, also called the epidermis, synthesize and secrete mucilage. At maturity, the mucilage dehydrates into a thin, compressed layer.

1.3.6 Accumulation of seed storage compounds

In *Arabidopsis*, the vascular bundle terminates at the end of the funiculus and there is no vascular tissue within the seed. The outer integument is thought to act as a symplastic extension of the phloem, allowing nutrients to move freely from cell-to-cell through plasmodesmata (Stadler *et al*., 2005). In contrast, the inner integument, endosperm and embryo represent separate symplastic domains; therefore the transfer of nutrients between these domains is thought to be apoplastic, requiring the activity of transporter proteins. Plants transport relatively simple compounds in the phloem, mainly sucrose and amino acids. Both sucrose (Baud *et al*., 2005) and amino acid (Schmidt *et al*., 2007) transporters have been associated with nutrient uptake into developing seeds.

In developing *Arabidopsis* seeds, the level of imported sucrose and amino acids decreases progressively, reflecting their utilisation in the formation of specialised storage compounds (Fait *et al*., 2006). The accumulation of carbohydrates, proteins and lipids is crucial for the survival of the next plant generation and is essential for both human and animal nutrition. Starch accumulates to high levels in developing *Arabidopsis* seeds, however it is only present transiently, reaching a peak at 10 DAP and almost completely disappearing in mature seeds (Focks and Benning, 1998; Baud *et al*., 2002). Although there have been many suggestions as to the purpose of this transient starch accumulation, its role in seed development is not yet clear (Baud *et al*., 2008). The synthesis and accumulation of proteins and lipids begins later than that of starch, and continues progressively until seed maturity. Developing *Arabidopsis* embryos accumulate lipids in the form of triacylglcerols (esters of glycerol and fatty acids) and the major seed storage proteins are 12S globulins and 2S albumins (Heath *et al*., 1986).

1.3.7 Comparison with seed development in crop species

Although seeds from different plant species appear to vary greatly in their size, shape and colour, the majority of their development follows the same basic pattern. One notable difference is that the endosperm of monocot species persists longer than that of eudicots (Lopes and Larkins, 1993). In cereals, the endosperm is retained as a reserve tissue in the mature seed and consists of four cell types: starchy endosperm, embryo-surrounding region, transfer cells and the aleurone layer (Olsen, 2004). The aleurone layer, which

covers the perimeter of the endosperm, produces enzymes that breakdown the storage reserves during seed germination. In contrast, the transient endosperm of *Arabidopsis* and *B. napus* is substantially degraded in mature seeds; however the peripheral cell layer remains and is often referred to as aleurone (Vaughn and Whitehouse, 1971; Groot and Van Caeseele, 1993). The role of aleurone cells in *Arabidopsis* is unknown as they contain few storage products and have thin cell walls (Keith *et al*., 1994). While the endosperm accumulates the storage reserves in monocots, the cotyledons act as the primary storage organ in eudicots. In *B. napus* and *Arabidopsis*, 90% and 60% respectively of seed oil was found to be stored in the cotyledons (Li *et al*., 2006).

1.4 Seed size regulation

Seeds are not simply containers, accepting what nutrients are supplied by the mother plant; they themselves have a partial role in determining the extent of their growth (Egli, 2006). The three seed components, the embryo, endosperm and seed coat, undergo co-ordinated growth through cell division and expansion to determine the final size of the seed (Garcia *et al*., 2005). Below, physiological and genetic factors that effect seed size are described, including many of the molecular regulators that have recently been identified.

1.4.1 Seed size versus seed number trade-off

Seed size is associated with seed number due to the limited resources of the mother plant (Harper *et al*., 1970; Venable, 1992). For example, a plant which allocates enough assimilate to reproduction to give a total seed yield of 500mg, could produce 10 seeds weighing 50mg each, or 5 seeds weighing 100mg each. Many examples of the negative correlation between seed size and number have been documented in *Arabidopsis* (Sills and Nienhuis, 1995; Alonso-Blanco *et al*., 1999; Jofuku *et al*., 2005; Schruff *et al*., 2006) and crop plants (Kiniry, 1988; Kiniry *et al*., 1990). While seed number appears greatly influenced by environmental conditions and resource availability, seed size is primarily controlled by genetic factors and as a result is more stable (Sadras, 2007).

1.4.2 Genetic regulation

Quantitative trait loci (QTL) mapping has been carried out in rice, tomato, soybean and sunflower, revealing many genetic loci associated with seed size. Recently, a few of these studies have led to the cloning and functional characterisation of genes involved in seed size regulation. The rice *GS3* QTL encodes a putative transmembrane protein associated with increased grain weight and length (Fan *et al*., 2006). Loss of function of the rice *GW2* QTL, which encodes a previously unknown RING-type E3 ubiquitin ligase, results in heavier grains due to increased cell numbers and seed filling and a 19.7% increase in yield (Song *et al*., 2007). Also in rice, an unknown gene within the *qSW5* QTL has been associated with increased grain size (Shomura *et al*., 2008). In tomato, an ABC transporter gene has been pinpointed as the cause of increased seed size associated with the *Sw4.1* QTL (Orsi and Tanksley, 2009). The authors also identified the *Arabidopsis* ABC transporter ortholog, and show that T-DNA mutants had altered seed development possibly as a result of altered lipid accumulation.

The seed mass of *Arabidopsis* ecotypes can vary up to 3.5-fold (Krannitz *et al*., 1991). Cvi seeds are reported to be almost double the weight of those produced by L*er*, due

to increases in the rate and duration of seed coat and endosperm development (Alonso-Blanco *et al*., 1999). A QTL analysis between these two ecotypes identified 11 loci affecting seed weight and length. A number of the seed size QTLs co-localised with QTL affecting ovule number, fruit size and total leaf number, highlighting the importance of these maternal factors in controlling seed size.

1.4.3 Endosperm development

The endosperm constitutes a large proportion of the mature seed in monocots hence seed size is directly attributed to endosperm proliferation (Cochrane and Duffus, 1983). The importance of endosperm proliferation in maize is emphasised by the *miniature1* (*mn1*) mutant which has dramatically reduced endosperm growth and small seeds (Cheng *et al*., 1996; Vilhar *et al*., 2002). However, endosperm proliferation has also been shown to have a strong, although indirect, effect on final seed size in the eudicot, *Arabidopsis*. Cell division in the endosperm is influenced by parent-of-origin effects (Scott *et al*., 1998). When a diploid seed parent is crossed with a tetraploid pollen parent the offspring have an excess of paternal genomes. Paternal excess results in over-proliferated endosperm, a delay in the onset of endosperm cellularisation and final seed size is increased compared to seeds produced in a balanced cross. In contrast, the reciprocal cross leads to maternal excess, reduces endosperm proliferation and small seeds are produced. These parent-of-origin effects on seed size are believed to occur due to genomic imprinting, the differential expression of genes from maternally and paternally derived genomes. The phenomenon of genomic imprinting is thought to have arisen as a consequence of conflict between the parents over resource allocation from the mother to the embryo (Haig and Westoby, 1989, 1991). Supporting this theory, seed phenotypes observed in interploidy crosses can be phenocopied by crossing plants with defects in DNA methylation, which is known to be an essential component of genomic imprinting (Adams *et al*., 2000).

Recently, three mutants have been identified in *Arabidopsis* with seed phenotypes similar to that found in seeds with an excess of maternal genomes. Early endosperm growth is restricted in *haiku1* (*iku1*), *iku2* and *miniseed3* (*mini3*) mutant seeds resulting in premature endosperm cellularisation, reduced expansion of cells in the seed coat and subsequently reduced seed size (Garcia *et al*., 2003; Luo *et al*., 2005). Due to the similar phenotype and gene expression profiles of these mutants the authors suggest that they may act in the same pathway (Luo *et al*., 2005). In support of this hypothesis, *MINI3* and *IKU2* expression is reduced in *iku1* plants and *IKU2* expression is reduced in *mini3* plants. *IKU* and *MINI3* genes are thought to regulate the onset of cellularisation or may directly control nuclear proliferation in the endosperm in order to influence seed size.

1.4.4 Integument and seed coat development

The growth and development of the seed coat, which derives from the ovule integuments, represents a means by which the mother plant can regulate seed size. Evidence suggests that the seed coat, which surrounds the embryo and endosperm, may physically restrict seed growth. *In vitro* grown soybean (*Glycine max*) seeds with the seed coat removed are able to grow larger and for longer than seeds developing *in vivo* (Egli, 1990). Small seeds are produced in the *Arabidopsis* mutant, *transparent testa glabra 2* (*ttg2*), as integument cell expansion fails (Garcia *et al*., 2005). Furthermore, large pointed seeds are produced in *auxin response factor 2* (*arf2*) mutants due to extra cell divisions in the ovule integuments and the subsequent formation of enlarged seed coats (Schruff *et al*., 2006). However, in plants containing the *p35S::KRP2* transgene, the constitutive over-expression of the *KIP RELATED PROTEIN 2* (*KRP2*) reduces the number of cells in the seed coat but seeds still reach wild-type proportions due to increased cell elongation (Garcia *et al*., 2005). Such compensation highlights that the seed coat does not have sole control of final seed size and suggests that cross talk exists between maternal and zygotic regulators.

Many *Arabidopsis* mutants with defective development of the ovule integuments give rise to small seeds with abnormal shape. Seeds of the *aberrant testa shape* (*ats*) mutant are small and heart-shaped due to the development of only a single integument consisting of three to five layers (Léon-Kloosterziel *et al*., 1994). The mutated gene responsible for the *ats* phenotype encodes a member of the KANADI family of transcription factors (McAbee *et al*., 2006). A similar seed phenotype was found in mutants of the *ARABIDOPSIS CRINKLY 4* (*ACR4*) gene which encodes a membrane-localised receptor-like kinase (Tanaka *et al*., 2002; Gifford *et al*., 2003). In *acr4* mutants, integuments cells lack organisation, form sporadic outgrowths and occasionally inappropriate cell types, such as stomata, occur. Integument development is severely reduced in *aintegumenta* (*ant*) mutants however large seeds can be obtained in plants over-expressing this transcription factor-encoding gene (Krizek, 1999; Mizukami and Fischer, 2000). Plants containing the *p35S::ANT* transgene show increased size of floral organs including seeds due to an extended period of cell division. Both the *INNER NO OUTER* (*INO*) and *SUPERMAN* (*SUP*) genes encode transcription factors (Sakai *et al*., 1995; Villanueva *et al*., 1999) and mutations lead to altered seed growth. While *ino* seeds are narrow and elongated (Baker *et al*., 1997), *sup* mutants produce seeds showing a range of shapes and sizes (Gaiser *et al*., 1995). Both mutants lack the normal curvature seen in wild-type ovules. In *ino* this is a due to a complete absence of the outer integument whereas in the *sup* mutant the outer integument is present but develops abnormally.

A further *Arabidopsis* gene that acts maternally to influence seed size is *APETALA 2* (*AP2*), which encodes a transcription factor involved in establishing the floral meristem, floral organ identity and development of the ovule and seed coat (Jofuku *et al*., 1994). Loss-of-function *ap2* mutants have defective seed coats with large, irregularly-shaped epidermal cells that lack mucilage. Seeds of *ap2* mutants are large, accumulate more storage products and contain larger embryos (Jofuku *et al*., 2005; Ohto *et al*., 2005). One study found that developing *ap2* seeds also had altered sugar composition compared to wild-type seeds (Ohto *et al*., 2005). Although the altered seed coat structure may in part be responsible for increased seed size in *ap2* mutants, an alternative hypothesis is that the high hexoses to sucrose ratio extends the period of cell division in the seed and increases seed growth. As many studies have suggested that sugar metabolism is an important mediator of seed size, this topic will be discussed in more detail below.

1.4.5 Metabolic control

Metabolic factors such as sugars can act as regulators, influencing gene expression and changing developmental programmes. Within developing seeds, changes in sugar composition are believed to control the shift from cell division to storage product accumulation (Weber *et al*., 1998). A high hexoses to sucrose ratio is positively correlated with cell division whereas a high sucrose level is associated with seed filling. As sugars are imported into the seed primarily as sucrose, the high hexoses to sucrose ratio is achieved by the activity of cell wall-bound invertases which hydrolyse the sucrose into glucose and fructose. Large-seeded genotypes of the broad bean *Vicia faba* have prolonged activity of seed coat-associated invertases, thus creating a high hexoses to sucrose ratio and promoting growth by increased cell division in the embryo (Weber *et al*., 1996). Although originally observed in legume species, similar shifts in sugar composition have been found in *Arabidopsis* (Focks and Benning, 1998) and *B. napus* (Hill *et al*., 2003). The small-seeded *mn1* mutant of maize has reduced endosperm growth due a mutation in a gene encoding a cell wall-bound invertase that accumulates in the endosperm transfer layer (Cheng *et al*., 1996). Sugar metabolism therefore has an important role in regulating seed development and size in a wide range of species.

1.5 Research aim and objectives

The overarching aim of the research presented here was to expand our understanding of seed development and seed size regulation in order to address the following question: can increasing seed size through the genetic modification of seed development produce a potentially useful increase in seed yield?

- Using previously reported information on seed development and seed size regulation, strategies to increase seed size and subsequently seed yield were designed and tested. These strategies included modifying the development of both the triploid endosperm and the ovule integuments/seed coat. Exploiting the model plant *Arabidopsis*, these experiments aimed to establish a proof-of-concept that could find application in economically-important seed crops.
- Seed yield can only be assessed objectively when fertility, and therefore seed number, are equivalent in the test line and wild-type plants. As *Arabidopsis* has an extremely large number of flowers, it would not be feasible to carry out manual pollinations in order to force full fertility. Therefore, partial yield may be assessed by restricting pollinations to allow only a certain number of flowers to set seed. However where the test line is fully fertile, yield may be assessed by comparing the total seed mass/plant. Crop plants are assessed by both their seed yield and yield efficiency, referred to as the HI, as this takes into account the total biomass of the crop. With the aim of providing a more realistic measurement of yield, the HI was utilized to assess the impact of traits, such as increased integument size, introduced into *Arabidopsis*.
- In order to further understand the molecular regulation of seed size, with particular reference to the role of the integuments/seed coat, a search was conducted for genes that could suppress or enhance the distinctive *arf2* seed phenotype, which is characterised by large pointed seeds due to extra cell division in the integuments. The goal was to identify genes that may interact with *ARF2* or operate independently to influence seed size. However, this project also prompted an in depth examination of how the development of the ovule integuments can determine seed shape.
- This thesis also contains a detailed study of the natural variation in seed size, yield and HI which aimed to elucidate important yield components in *Arabidopsis*. Results of this analysis confirmed the trade-off between seed size and number and prompted a new investigation to determine the physiological and developmental basis of this yield component compensation.

2. MATERIALS AND METHODS

Unless otherwise stated chemicals and products were sourced from Sigma.

2.1 MATERIALS

2.1.1 Plants

Seed stocks of Col-0, L*ER*, L*er*, Col-*er*, Ob-0, Rsch-0 and Bla-1 ecotypes were obtained from the Nottingham Arabidopsis Stock Centre (NASC). The *mnt*/*arf2-9* mutant was discovered in a population of EMS-mutagenised Col-3 seeds that were supplied alongside wild-type Col-3 seeds (Lehle Seeds). The Salk_108995/*arf2-8* (Col-0 background) T-DNA insertion line (Alonso *et al*., 2003) was sourced from the SALK database (signal.salk.edu/cgi-bin/tdnaexpress) and obtained from NASC. The *gai* mutant (L*er* background) was also obtained from NASC. Kas-1 seeds were kindly donated by T.J. de Jong (Leiden University, The Netherlands). Cvi seeds were kindly donated by M. Kornneef (Max Planck Institute for Plant Breeding Research, Germany). The *acr4-2* mutant (Col-0 background) was kindly donated by G. Ingram (University of Edinburgh, UK). *ap2-11* seeds (Col-0 background) were kindly donated by J.J. Harada (University of California, US). *pAP1::ARF2* and *pINO::ARF2 RNAi* transgenic lines were provided by M.C. Schruff (University of Bath, UK).

2.1.2 Bacteria

Two *Escherichia coli* strains were used for cloning: Subcloning Efficiency DH5α™ chemically-competent cells (Invitrogen) and TransforMax™ EC100™ electrocompetent cells (Epicentre Biotechnologies). Plant transformation was carried out using the *Agrobacterium tumefaciens* strain GV3101 (Koncz and Schell, 1986) which harbours the non-oncogenic Ti plasmid (pGV3101).

2.1.3 Bacterial growth media

E. coli was grown in LB liquid media containing 5 g/L NaCl, 5 g/L yeast extract (Fisher Scientific) and 10 g/L tryptone (Fisher Scientific), with the pH adjusted to 7.2 with NaOH. *A. tumefaciens* was grown in 2YT media containing 5 g/L NaCl, 5 g/L yeast extract and 16 g/L tryptone, with the pH adjusted to 7.2 with NaOH. For solid media, 15 g/L granulated agar (Difco) was added.

2.1.4 Plasmids

Plasmids used for cloning include: the pGEM®-T vector (Promega), BJ36, BJ40 and BJ60 (kindly donated by Bart Janssen, Horticultural & Food Research Institute, New Zealand) and the RNAi vector, pFGC5941 (Arabidopsis Biological Resource Centre (ABRC)). pGEM®-T, BJ36 and BJ60 confer carbenicillin resistance in bacteria. BJ40 confers spectinomycin resistance in bacteria. Both BJ40 and BJ60 confer kanamycin resistance in plants. The RNAi vector, pFGC5941 confers kanamycin resistance in bacteria but basta resistance in plants.

2.1.5 Primer design and synthesis

NetPrimer (Premier Biosoft International) was used to check all primers and determine melting temperatures. After primer design, a megaBLAST (NCBI) search for highly related nucleotide sequences across the entire *Arabidopsis* genome was carried out to ensure target specificity. Primers were ordered from either Sigma-Genosys or Invitrogen. Primer information is given in Table 2.1 and Table 2.3.

2.1.6 Sequencing services and data analyses

Sequencing was carried out by Cogenics. Chromatograms were viewed using Chromas Lite version 2.01 (Technelysium Pty Ltd) and sequence alignments were carried out using GeneDoc version 2.6.02 (www.nrbsc.org/gfx/genedoc/index.html). The cloning of molecules was planned and visualised using Clone Manager 5 (Sci Ed Central).

2.1.7 Statistical analyses

Statistical tests were performed using Minitab 14 software. Comparisons for normally distributed data with equal variances were carried out using Student's *t*-test or one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons (family error rate, 5). Where appropriate, the raw data were transformed (log10) in order to give a normal distribution and uniform variance. Data not following a normal distribution was analyzed using the Mann-Whitney *U*-test with the P-value adjusted for ties.

2.1.8 Image capture and processing

Photographs of whole plants and inflorescences were obtained using a Coolpix 4500 digital camera (Nikon). Mature seeds were photographed under a SMZ1500 dissecting microscope (Nikon) using a Digital Sight DS-U1 colour camera (Nikon). Images were processed using Adobe® Photoshop® Elements.

2.2 METHODS

2.2.1 Plant growth conditions

Seeds were stratified for 3–5 days at 4 ˚C in 0.15% agar. Stratified seeds were transferred onto Levingtons F2 compost pre-treated by soil drench with 0.02% Intercept (Scotts) and covered with a propagator lid until 10 days after germination. Plants were grown in a Sanyo controlled environment room with a daylength of 16 h, 70% relative humidity, 22 ˚C (day) and 18 ˚C (night), or in a glasshouse at 24 ˚C (day) and 17.5 ˚C (night).

For total seed yield and HI analysis, plants were grown in 3 inch pots and fitted with an Aracon (Lehle Seeds) when the primary inflorescence was 10 cm high (Figure 2.1a). If plants grew taller than the Aracon, a further plastic tube was placed on top of the first to ensure that all seeds were collected.

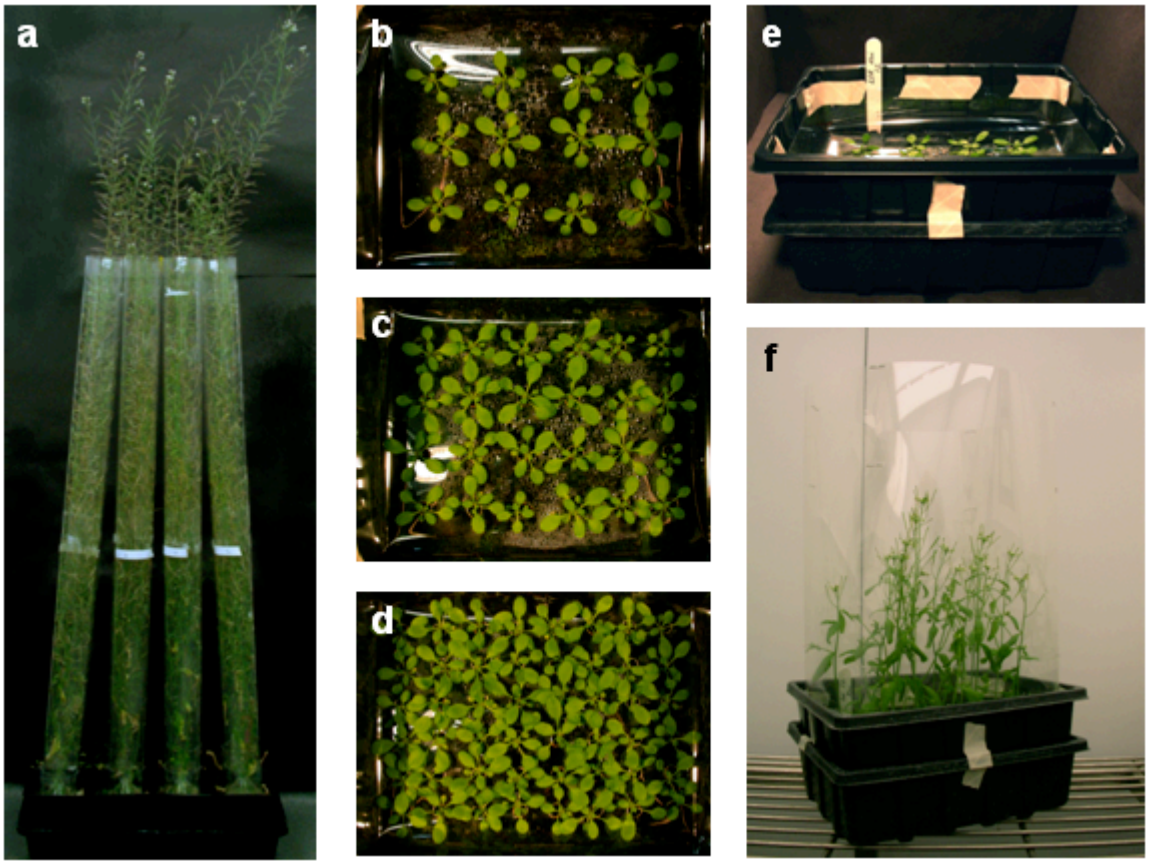

Figure 2.1: Growing *Arabidopsis* for yield, HI and density experiments.

(a) Plants grown for seed yield and HI analyses were fitted with Aracons to prevent loss of seed. Density experiments were carried out at (b) 375, (c) 750 and (d) 1500 plants/m^2. (e) Seeds germinate through holes in plastic sheet and develop into rosettes. (f) Flowering plants are surrounded by plastic to prevent seed loss.

To determine how density may affect seed size, yield and HI, three different densities were tested: 12, 24 and 48 plants per half tray (16 cm x 20 cm). This equated to 375, 750 and 1,500 plants/m^2 respectively (Figure 2.1b,c,d). Plants were grown in half trays containing 700 g Levingtons F2 compost. A plastic sheet with evenly-spaced, 1 cm diameter, holes was placed inside a further half tray with the base removed. This structure was placed directly onto the soil and a seed placed in the centre of each hole (Figure 2.1e). When the plants started flowering, the plants were surrounded with plastic sheeting to prevent seed loss (Figure 2.1f).

For plant transformations, 15 Col-0 seeds were planted in each of three 6 inch pots per construct. Plants were grown for 4-6 weeks until the primary inflorescences were 10 cm high and then chopped to encourage the growth of multiple secondary inflorescences. Approximately 4-5 days later plants should have sufficient buds and open flowers for floral dipping.

For testing kanamycin resistance, seeds were surface-sterilised as follows: 5 min in 70% ethanol, 5 min in 50% bleach with 0.05% Tween 20, followed by 6 washes in sterile distilled H_2O. After sterilisation, seeds were transferred to Petri dishes containing growth media (full strength Murashige and Skoog Media with Gamborg's Vitamins (Sigma), 1% sucrose and 0.8% Phyto Agar (Duchefa), pH adjusted to 5.8 with 0.1 M KOH) and kanamycin (50 µg/ml). Plates were sealed with parafilm and seed was stratified for 3-5 days at 4 ˚C. Plates were then subjected to 6 h light treatment, followed by 48 h dark treatment, and then transferred to standard conditions in the Sanyo controlled environment room (detailed above) to facilitate distinguishing resistant and non-resistant seedlings (Harrison *et al*., 2006). After approximately two weeks, resistant seedlings were transplanted into Levingtons F2 compost.

For testing BASTA (a commercial formulation of glufosinate) resistance, seeds were sown directly on to Levingtons F2 compost. Following germination, seedlings were sprayed with 427.5 µl/L BASTA (AgrEvo) until completely covered on two separate occasions.

2.2.2 Cross pollinations

One day prior to anthesis, anthers were dissected from floral buds of the seed parent. A few flowers from above and below the emasculated buds were removed so as not to confuse self- and cross-pollinated siliques. Two or three days following emasculation, when stigmatic papillae are fully expanded, mature pollen from the pollen parent was gently dabbed onto the mature stigmas. Both emasculation and pollination were carried out under a Leica MZ6 dissecting microscope.

2.2.3 Restricted and unrestricted pollinations

In restricted pollinations, six siliques on the primary inflorescence were allowed to set seed and all other floral buds on the plant were removed (Figure 2.2a). As *arf2-9*, *pAP1::ARF2* and *ap2-11* plants show reduced fertility, full seed set in each of the six siliques was ensured by manual pollination. In unrestricted pollinations, all flowers were allowed to self-pollinate (Figure 2.2b).

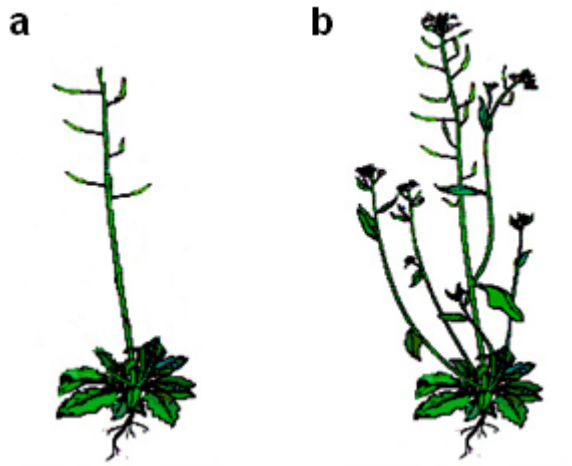

Figure 2.2: Diagrammatic representation of (a) restricted and (b) unrestricted pollinations.

2.2.4 Mature seed weight, total seed yield and HI analyses

Mature seeds were separated from dry aerial organs using a sieve and stored in 1.5 ml tubes with pierced lids. Occasionally, seed was stored in a sealed container with silica gel to accelerate drying. All aerial organs including rosette leaves, stems and empty silique cases were left for 1 week to dry and stored in a sealed jar with silica gel for 24 h before weighing. Mean seed weight was determined by weighing batches of 50 seeds using an UMT2 microbalance (Mettler-Toledo). Total seed yield and the dry mass of aerial organs were determined using a Sartorius A200S electronic balance (Sartorius Analytic). HI was calculated as the ratio of total seed yield to biological yield (total seed yield plus the dry mass of aerial organs).

2.2.5 Fresh and dry weight analyses of developing seeds

Developing seeds were dissected from siliques and fresh weight was determined immediately for 2 batches of 10 seeds per silique. Seeds were then transferred to 50 ˚C for 2 days before dry weight was also assessed by weighing 2 batches of 10 seeds per silique. Water content was calculated as fresh weight minus dry weight.

2.2.6 Plant DNA extraction

Method adapted from Edwards *et al*., (1991) in order to extract DNA for PCR-based genotyping:

Two small, young leaves or an equivalent volume of another tissue, were placed in a 1.5 ml tube on ice with a toothpick measure of glass beads and 450 µl extraction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl and 25 mM EDTA). The leaf material was ground using a bench drill and sterile plastic pestle and centrifuged at maximum speed in a bench-top centrifuge for 15 min. The supernatant was transferred into a sterile 1.5 ml tube and centrifuged again. 350 µl supernatant was transferred into a sterile 1.5 ml tube and an equal volume of 100% isopropanol was added, followed by incubation at room temperature for 5-30 min. After a further centrifugation, the pellet was washed three times with 70% ethanol, then left to dry in a laminar flow hood for 5 min. The pellet was dissolved in 80 µl sterile milli-Q H_2O and the sample stored at 4 ˚C.

2.2.7 Plant RNA extraction

Fresh plant material was snap frozen in liquid nitrogen and if not used immediately stored at -80 ºC. The frozen tissue was ground using a ceramic pestle and mortar and RNA was isolated using the SV Total RNA Isolation System (Promega) according to the manufacturer's instructions.

2.2.8 cDNA synthesis

cDNA was synthesised from 1 µg total RNA using the ImPromII™ Reverse Transcription System (Promega) according to the manufacturer's instructions. PCR was then carried out using the cDNA as a template.

2.2.9 Amplification of DNA for genotyping and expression analyses

PCR was carried out using *Taq* DNA polymerase. Each 25 µl reaction contained 0.25 µl *Taq* (5 U/µl), 2.5 µl 10x PCR buffer (100 mM Tris-HCl pH8.3, 500 mM KCl, 15 mM $MgCl_2$), 0.5 µl dNTPs mix (10 mM), 1 µl of each primer (10 µM), 2 µl of DNA extraction (~10 ng) and 17.75 µl sterile milliQ H_2O. All reactions were carried out on a PTC-200 Thermal Cycler (MJ Research). A typical amplification programme consisted of an initial denaturation at 94 ˚C for 3.5 min, followed by 32 cycles of denaturation at 94 ˚C for 30 s, annealing at ~55 ˚C (depending on primer Tm) for 30 s and extension at 72 ˚C for ~1 min (depending on fragment length, 1 min/kb), and then a final extension at 72 ˚C for 10 min. Specific PCR conditions are given in Table 2.2.

Table 2.1: Primer sequences used for genotyping and expression analyses

Primer	Sequence (5'→ 3')	Tm (ºC)
0.05F	GAGTGGGTGGAGTGTGTTTG	55.4
0.05R	AGTTGGTTTTCGTTTGAGCAT	56.8
AGL28 R Check	GCAGTGTTGGAGCAGAAGTAG	55.1
ANT R Check	CAAGTTGAGGTGGAACAGAAG	54.7
AP1 F Check	ACCCAAATCCTAAAGAAACCACT	58.4
ARF2 R Check	CAGTTCCCTTCTCTTGACTTTG	56.1
ARF2 RT-PCR For	TCCTGAGGCTAATCAAGACGAG	58.9
ARF2 RT-PCR Rev	TCCTTGTTGTCGCATCGC	58.1
AVA ocs3' 2 R	TTGGATTTGTAATAATAAAACGCAAT	59.6
CZE F Check	ATTGTCTATGAGATGAGTAAGGGC	56.8
GAPC Forward	CACTTGAAGGGTGGTGCCAAG	62.7
GAPC Reverse	CCTGTTGTCGCCAACGAAGTC	62.8
INOMNTRNAi F	TTCTCTTGACTTTGCTGTCTCCT	57.9
INOMNTRNAi R	ATCTCCTATCATTCATCGTACACAC	57.2
lba1	TGGTTCACGTAGTGGGCCATCG	67.4
MNT R Check	GAAGTTATCTCCTCCACGATTACC	58.5
PER F Check	CACAATGGAATGGAGAGAGGAG	58.7
pFGC5941 ocs R2	TAGATGTCGCTATAAACCTATTCAG	55.7
PHE1 R Check	GACATCACTTCTTCAACGCCTT	58.7
SHP2 F Check	GCCGTTGGACTAACCTAATGT	56.3
SUGAR R Check	GAGGATTTGATTGTCGTATTTG	54.5
SUGAR RT-PCR For	AGCGTGGAGATGTTCTGTTAC	54.4
SUGAR RT-PCR For2	CTTGGAGTCTCTGGTGGTGTTAC	58.1
SUGAR RT-PCR Rev	GAAACGGATTCTTCACGGC	57.8

Table 2.2: PCR conditions used for genotyping and expression analyses

PCR	Primers	Annealing temp (˚C)	Extension time (s)	Fragment length (bp)
ARF2 transcript expression	ARF2 RT-PCR For, ARF2 RT-PCR Rev	58	30	400
GAPC expression	GAPC Forward, GAPC Reverse	56	40	543
pAP1::ARF2	AP1 F Check, MNT R Check	54	30	347
pCZE::SUGAR	CZE F Check, SUGAR R Check	54	90	1435
pINO::ARF2 RNAi	INOMNTRNAi F, INOMNTRNAi R	52	60	520
pPER::AGL28	PER F Check, AGL28 R Check	55	80	1318
pPER::ANT	PER F Check, ANT R Check	54	50	859
pPER::PHE1	PER F Check, PHE1 R Check	58	50	887
pSHP2::ARF2 RNAi(1)	SHP2 F Check, ARF2 R Check	56	105	1641
pSHP2::ARF2 RNAi(2)	INOMNTRNAi F, AVA ocs3' 2 R	56	50	839
Salk_108995 (genomic)	0.05F, 0.05R	54	150	1980
Salk_108995 (T-DNA)	0.05F, lba1	55	150	2124
SUGAR transgene expression	SUGAR RT-PCR For, pFGC5941 ocs R2	55	45	748
SUGAR endogenous expression	SUGAR RT-PCR For2, SUGAR RT-PCR Rev	55	40	696

2.2.10 Amplification of DNA for cloning and sequencing

The proof-reading KOD HiFi DNA polymerase (Novagen) was used to amplify DNA fragments for cloning and sequencing. Each 50 µl reaction contained 2 µl DNA (~10 ng), 0.5 µl KOD (2.5 U/µl), 5 µl 10x buffer#1 (supplied with KOD), 2 µl $MgCl_2$ (25 mM), 5 µl dNTPs (2 mM), 2 µl of each primer (10 µM) and 31.5 µl sterile milliQ H_2O. All reactions were carried out on a PTC-200 Thermal Cycler (MJ Research). The PCR programme was the same as for *Taq* DNA polymerase except the final extension was 5 min. Specific PCR programmes are given in Table 2.4.

Table 2.3: Primer sequences used for cloning and sequencing (proof-reading KOD polymerase)

Primer	Sequence (5'→ 3')	Restriction Site	Tm (ºC)
ACR4 F1	TGAAGACAGTAACTAAGAAGCGTG	-	56.7
ACR4 R1	CAACCCGTCAAGAATGCCAC	-	61.0
ACR4 F2	CTTGTTGATTGTTGGGGTTAC	-	54.8
ACR4 R2	CTTTCCTTTTTCCTTGCCTC	-	56.2
ACR4 F3	GTCTTGGTTTCCCTGCTTCTATC	-	59.2
ACR4 R3	GCCTTTCTCCACATTCTTCAC	-	56.4
ACR4 F4	CTTGAAAAAGCCGCTGATGG	-	60.9
ACR4 R4	ATTCTACTATGTTCCCTTCTTCG	-	54.8
ACR4 F5	GATTTTGGTCTCTCCTTACTTGG	-	57.1
ACR4 R5	GAGTCTCTCTGCCTCTTTGTTAC	-	54.8
At1g01530 AGL28 cds F	ttt*ggtacc*ATGGCGAGAAAGAATCTTGG	*Kpn1*	57.0
At1g01530 ALG28 cds R	aaa*cccggg*CTAATAGTAACGAGCCCAATAC	*Xma1*	51.7
At1g65330 PHE1 cds F	aaa*ggtacc*ATGAGGGGGAAGATGAAG	*Kpn1*	51.3
At1g65330 PHE1 cds R	ttt*cccggg*CTAGAGATCATTGATGATGTTAGG	*Xma1*	53.4
At1g77210 SUGAR cds F	ttt*ggtacc*ATGGCCGGTGGAGCTCTTACCG	*Kpn1*	69.2
At1g77210 SUGAR cds R	ttt*cccggg*TTATTCATCAACATCTTCGACATATTTC	*Xma1*	60.8
At2g42830 SHP2 prom F	aaa*gaattc*AAAAAAAAATATCTCAAAGTAAACG	*EcoR1*	53.3
At2g42830 SHP2 prom R	aaa*ggcgcgcc*TTCTATAAGCCCTAGCTGAAGTAT	*Asc1*	54.7
At4g37750 ANT cds F	ttt*cccggg*TCAATCCAGCACAATGATGTTATCC	*Kpn1*	63.0
At4g37750 ANT cds R	aaa*cccggg*TCAAGAATCAGCCCAAGCAGCG	*Xma1*	67.2
At5g07210 CZE prom F	ttt*gtcgac*ATGGGATGATCTCCGTTACC	*Sal1*	56.1
At5g07210 CZE prom R	ttt*ggtacc*TCTAATAATCTTTGCAAAGAG	*Kpn1*	47.7
At5g46950 PER prom F	ttt*gtcgac*AATTGATGGAAATAAAATTTCGC	*Sal1*	58.0
At5g46950 PER prom R	ttt*ggtacc*TTTTTTGTTTTTACTTTGAGAAGAAG	*Kpn1*	57.0

Linkers are in lowercase letters, restriction sites are in italics. When linkers are added the Tm does not include the linker.

Table 2.4: PCR conditions used for cloning and sequencing (proof-reading KOD polymerase)

PCR	Primers	Annealing temp (˚C)	Extension time (s)	Fragment length (bp)
ACR4 Fragment 1	ACR4 F1, ACR4 R1	58	50	724
ACR4 Fragment 2	ACR4 F2, ACR4 R2	56	50	731
ACR4 Fragment 3	ACR4 F3, ACR4 R3	58	50	774
ACR4 Fragment 4	ACR4 F4, ACR4 R4	55	40	676
ACR4 Fragment 5	ACR4 F5, ACR4 R5	56	60	827
AGL28 cds	At1g01530 AGL28 cds F, AGL28 cds R	51	50	744
ANT cds	At4g37750 ANT cds F, ANT cds R	55	100	1668
PER promoter	At5g46950 PER prom F, PER prom R	60	100	2084
CZE promoter	At5g07210 CZE prom F, CZE prom R	57	100	1952
PHE1 cds	At 1g65330 PHE1 cds F, PHE1 cds R	51	50	840
SHP2 promoter	At2g42830 SHP2 prom F, SHP2 prom R	53	110	2100
SUGAR cds	At1g77210 SUGAR cds F, SUGAR cds R	60	90	1515

2.2.11 Colony PCR

Colony PCR was carried out as with standard *Taq* DNA polymerase except that the DNA template was replaced with 10 µl sterile milliQ H_2O, into which a sample of colony to be screened was transferred using a pipette tip. After mixing gently, the cells were lysed at 95 ˚C for 10 min and then PCR was performed as normal.

2.2.12 Agarose gel electrophoresis

0.8-1.5% agarose gels were made by dissolving agarose (Invitrogen) in 1x TAE buffer (40 mM Tris, 1 mM EDTA, 1.14% glacial acetic acid, pH 7.6) and before setting ethidium bromide was added to a final concentration of 0.2 µg/ml. Gels were left to set for at least 1 h, then placed in a Sub-Cell tank (Bio-Rad) and submersed in 1x TAE buffer containing ethidium bromide. Prior to loading of samples, the appropriate volume of 6x loading dye was added to each. Either a 1 kb or 100 bp DNA ladder (New England Biolabs) was loaded on each gel as a molecular weight marker. Electrophoresis was normally carried out at ~100 V using a Powerpac 300 power supply (Bio-Rad). DNA bands were visualized on a UV transilluminator (UVP).

2.2.13 Gel purification of DNA fragments

Following gel electrophoresis, PCR products were recovered using the Wizard® SV Gel and PCR Clean-Up System (Promega) according to the manufacturer's instructions.

2.2.14 A-tailing and plasmid ligation

Before ligation of DNA fragments into the pGEM®-T vector, PCR products were A-tailed. 10 µl reactions consisted of 7 µl gel purified DNA fragment, 1 µl 10x PCR buffer, 1µl dATP (0.2 mM) and 1 µl *Taq* DNA polymerase. The reaction mixture was then heated for 15 min at 70 ºC in a PTC-200 Peltier Thermal Cycler (MJ Research/Biorad).

Ligation reactions were carried out using T4 DNA ligase (Promega) according to the manufacturer's instructions. 12 µl reactions were incubated overnight at 4 ˚C. Usually a 3:1 molar ratio of vector to insert was used. 1 µl of ligation product was then transformed into either *E. coli* or *A. tumefaciens* cells.

2.2.15 DNA plasmid purification

A single bacterial colony was used to inoculate 5 ml LB containing the appropriate antibiotic. The culture was grown overnight at 37 ºC with shaking. Plasmid isolation was then carried out using the Wizard® Plus SV Minipreps DNA purification System (Promega) according to the manufacturer's instructions.

2.2.16 Restriction digests

Restriction digests were carried out in order linearise plasmids, cut genes from plasmids and to check that genes had been correctly inserted into plasmids. Restriction enzymes were used according to the manufacturer's instructions (*BamH*I, *EcoR*I, *Kpn*I, *Not*I, *Sal*I, *Xba*I and *Xma*I supplied by Promega; *Asc*I and *Swa*I supplied by New England Biolabs). All reactions were incubated overnight at 37 ºC. Plasmid fragments were separated by gel electrophoresis and those required for cloning were gel purified.

2.2.17 Preparation of chemically-competent E. coli

A 10 ml LB overnight culture of Subcloning Efficiency DH5α™ chemically-competent cells (Invitrogen) was set up. The following day, 250 µl of culture was transferred to 50 ml LB and shaken at 37 ºC until the OD_{600} = 0.437. The cells were left on ice for 30 min, centrifuged at 1,388 x g at 4 ºC for 10 min and then resuspended in 50 ml ice-cold 0.1 M $CaCl_2$. The incubation and centrifugation steps were repeated, and the cells resuspended in 8 ml ice-cold 0.1 M $CaCl_2$. After a further 30 min incubation on ice, 2 ml sterile, ice-cold 100% glycerol was added. 50 µl aliquots were dispensed into pre-chilled tubes, snap-frozen in liquid nitrogen and stored at -80 ºC.

2.2.18 Preparation of electrocompetent A. tumefaciens

A 10 ml 2YT culture of *A. tumefaciens* GV3101 containing rifampicin (100 µg/ml), gentamycin (25 µg/ml) and carbenicillin (100 µg/ml) was grown overnight at 28 ºC. 250 µl of starter culture was used to inoculate a 50 ml culture which was again grown overnight at 28 ºC. Cells were cooled on ice, centrifuged at 4,615 x g for 6 min at 4 ºC. The pellet was resuspended in 50 ml sterile, ice-cold 10% glycerol. Centrifugation and resuspension was repeated with 20 ml, 5 ml, 1 ml and 200 µl glycerol. 40 µl aliquots were dispensed into pre-chilled tubes, snap-frozen in liquid nitrogen and stored at -80 ºC.

2.2.19 Transformation of chemically-competent E. coli

50 µl of chemically competent *E. coli* were thawed on ice and 1-2 µl of plasmid or ligation mix was added and mixed gently. After a 30 min incubation on ice, the cells were heat-shocked for 20 s in a 37 ºC water bath, and returned to ice for 2 min. 750 µl of pre-warmed LB was added and the cells shaken at 225 rpm at 37 ºC for 1 h. 100 µl and 200 µl of transformation mix was spread on LB plates containing the appropriate antibiotics (100 µg/ml carbenicillin or 25 µg/ml kanamycin). The remaining transformation mix was stored at 4 ºC and plated out the following day if required. Plates were incubated overnight at 37 ºC. When transforming plasmids with a pGEM®-T backbone, positive colonies were identified using blue/white selection on LB plates containing 10 µg/ml IPTg and 50 µg/ml X-gal as well as antibiotic. Colonies were then screened for presence of the correct plasmid using colony PCR.

2.2.20 Transformation of electrocompetent E. coli and A. tumefaciens

30-50 µl of electrocompetent cells were thawed on ice and 1-2 µl of plasmid or ligation mix was added and mixed gently. The mixture was transferred to a pre-cooled 0.2 cm electroporation cuvette (Molecular BioProducts). Electroporation was performed using a MicroPulser™ (Bio-Rad) on setting Ec2 (2.5 kV, 1 pulse). For transforming *E. coli*, immediately following electroporation, 1 ml pre-warmed LB was added and the cells shaken at 37 ºC for 1 h. 100 µl and 200 µl of transformation mix was spread on LB plates containing the appropriate antibiotics (100 µg/ml carbenicillin or 25 µg/ml kanamycin). LB plates were incubated overnight at 37 ºC. For transforming *A. tumefaciens*, 1 ml pre-warmed 2YT was added and the cells shaken at 28 ºC for 2 h. 100 µl and 200 µl of transformation mix was spread on 2YT plates containing 100 µg/ml rifampicin, 25 µg/ml gentamycin and 100 µg/ml carbenicillin. 2YT plates were incubated at 28 ºC for approximately 40 h.

2.2.21 Transformation of Arabidopsis thaliana

Protocol adapted from Clough and Bent (1998). A starter culture of 50 ml 2YT plus antibiotics (100 µg/ml rifampicin, 25 µg/ml gentamycin and 100 µg/ml carbenicillin) was inoculated with a streak of *A. tumefaciens* cells using a sterile loop and shaken (200 rpm) overnight at 28 ºC. 10 ml of the starter culture was used to inoculate a 500 ml culture, which was shaken overnight at 28 ºC. Bacterial cells were harvested by centrifugation at 4,615 x g for 15 min at room temperature. The pellet was resuspended in 500 ml 5% sucrose. Silwett L-77 (Lehle Seeds) was added to a final concentration of 0.036%. The solution was poured into a tilted seed tray and *Arabidopsis* flowers submerged for 10 s

with agitation. After excess solution had been removed from the inflorescences, plants were covered in an autoclave bag and stored overnight in shade. The following day, plants were uncovered and returned to their normal growth environment. Seeds were collected at maturity and transformants selected using either kanamycin or basta.

2.2.22 Selection of homozygous single insertion lines

100 T_2 generation seeds from each independent positive transformation line were screened for either kanamycin or basta resistance. Lines demonstrating approximately 75% resistant seedlings were assumed to carry a single copy of the transgene (single insertion line). Several resistant plants from this screen were allowed to set seed. T_3 generation seed was then subjected to kanamycin or basta and lines demonstrating 100% resistance were assumed to be homozygous for the transgene. Single insertion homozygous T_3 seed was used for all further investigation.

2.2.23 Seed clearing and differential contrast (DIC) microscopy

To visualise GUS activity and measure embryo sac area and integument cell number, developing seeds were dissected from siliques and mounted in clearing solution (chloral hydrate/water/glycerol, 8w/3v/1v). Cleared seeds were imaged using differential contrast optics on a 90i eclipse microscope (Nikon) and photographed with either a Digital Sight DS-U1 colour camera (Nikon) or a Digital Sight DS-1QM/H black and white camera (Nikon).

2.2.24 Histochemical staining for localisation of GUS activity

Developing seeds were dissected from siliques and submerged in GUS staining buffer (100 mM potassium phosphate buffer pH 7.0, 1 mg/ml X-Gluc, 0.1% Triton X-100, 1-3 mM $K_3Fe(CN)_6$ and 1-3 mM $K_4Fe(CN)_6$) in a 96-well plate. The plate was wrapped in foil and incubated overnight at 37 ºC. Seeds were cleared and viewed using DIC microscopy.

2.2.25 Feulgen staining and confocal microscopy

Protocol adapted from Braselton *et al*., (1996).

FIXATION: *Arabidopsis* siliques (3-6 DAP) were trimmed at both ends using a razor blade, placed in 1 ml of fixative (3:1 ethanol:acetic acid) and stored overnight at 4 ºC. Siliques were transferred to 70% ethanol and stored at 4 ºC, where they could be kept for several weeks before use.

DAY 1: fixed siliques were washed three times in sterile milliQ H_2O for 15 min. After hydrolysis in 5 N HCl for 1 h, a further three washes in sterile milliQ H_2O for 10 min were

carried out. Siliques were transferred to Schiff's Reagent and stained for 2-3 h in the dark. The siliques were washed three times in cold tap water for 10 min. Following this, the siliques were washed once in 70% ethanol, once in 95% ethanol and twice in 100% ethanol, each for 10 min and then stored at 4 ºC overnight.

DAY 2: the 100% ethanol was changed every hour until the solution stopped turning pink. At this point, samples could be stored for several weeks at -20 ºC. The ethanol was replaced with 1:1 ethanol:LR White™ Resin (Agar Scientific) for 1 h and shaken occasionally. The siliques were placed in pure resin for 1 h and fresh resin was added before leaving overnight at 4 ºC.

DAY 3: a droplet of resin containing LR White™ Accelerator (Agar Scientific) was placed in the centre of a microscope slide and two coverslips were pushed partway into it from either side. Under a Leica MZ6 dissecting microscope, seeds were gently dissected out of the silique and the pod wall removed. A third coverslip was placed on top so that its edges were supported by the other two coverslips. After polymerisation of the resin, the slides were placed at 4 ºC where they may be kept for several months.

MICROSCOPY: Feulgen stained seeds were visualised on a 90i Eclipse microscope (Nikon). Florescence was detected using a 488 nm argon laser and a 515-530 nm filter.

2.2.26 Mutagenesis and screening

A preliminary EMS mutagenesis was at first carried out based on Jander *et al*., (2003). Four batches of 1,000 Salk_108995/*arf2-8* seeds were placed into 1.5 ml tubes and imbibed overnight in 1 ml sterile milliQ H_2O at room temperature. The seeds were exposed to 0.0%, 0.1%, 0.2% or 0.3% (w/v) EMS in 8.75 ml H_2O in 50 ml Falcon tubes on a rocking platform (speed = 45 rpm) at room temperature for 16 h. Tubes were laid at 15˚ from horizontal ensuring that the seeds were mixed thoroughly but not too vigorously in order to avoid seed damage. To inactivate the EMS, seeds were washed twice with 10 ml 0.1 M sodium thiosulphate for 15 min each time. After transferring to fresh tubes, the seeds were rinsed 10 times with H_2O and stratified at 4 ˚C for 3 days in 20 ml 0.15% agar. 100 seeds of each treatment were then germinated on filter paper to assess the germination rate. The remaining mutagenised seeds were sowed on to soil. The optimum dose of EMS was determined by calculating the percentage of embryo lethals within the siliques of M_1 generation plants. Siliques from 30 individual plants were examined for each EMS treatment.

The protocol described above was adapted to carry out a full scale mutagenesis using 5,000 Salk_108995/*arf2-8* seeds. All seeds were exposed to 0.2% EMS treatment and sowed onto

soil giving rise to the M_1 generation plants. The seeds from 100 M_1 plants were pooled to form an M_2 family. 800 seeds from each M_2 family were sown to produce M_2 plants that were then screened for flower opening, fertility and seed size and shape. Fertility was determined by recording the number of siliques that had elongated from the first 20 flowers on the primary inflorescence. To assess seed size and shape, the seeds were collected from individual M_2 plants and compared to non-mutagenised Salk_108995/*arf2-8* seeds under a Leica MZ6 dissecting microscope.

2.2.27 Determination of seed fatty acid content

FA analysis was carried out by Tony Larson and Ian Graham (University of York, UK).

FAs were measured in batches of 15 seeds by a one-step direct transmethylation procedure (Browse *et al*., 1986). Transmethylation was carried out for 2 h at 85 ºC using tripentadecanoin as an internal standard; this procedure has been demonstrated to be exhaustively quantitative for *Arabidopsis* (Li *et al*., 2006). The FAMEs formed were dissolved in hexane, and 1 µl aliquots were injected for gas chromatography-flame ionization detector analysis using a BPX70 10 m × 0.10 mm i.d. × 0.20 µm film thickness capillary column (SGE) and a GC-Ultra gas chromatograph (ThermoFinnigan). Injection was made into a hydrogen carrier gas stream at 0.3 ml/min at a 250:1 split ratio. After 0.1 min, the flow was increased to 0.5 ml/min. The temperature was ramped as follows: 150 ºC isothermal for 0.1 min; 16 ºC/min to 220 ºC; cooling at 120 ºC/min to 150 ºC; total analysis time, 5.1 min. FAMEs were identified by comparison with a Supelco 37 FAME mix and response factors relative to the internal standard were used to quantify individual FAs in samples using ChromQuest 4.2 software (ThermoElectron).

3. ENGINEERING ENLARGED SEED SIZE USING ENDOSPERM-LED APPROACHES

3.1 Introduction

The extent of endosperm proliferation is a major determinant of seed size and viability, both in the persistent endosperms found in cereals (Cochrane and Duffus, 1983) and the transient endosperms found in dicotyledonous species, such as *Arabidopsis* (Scott *et al*., 1998; Garcia *et al*., 2003; Luo *et al*., 2005). Since large seeds often have a higher endosperm cell number than small seeds, it is possible that increased endosperm proliferation enhances the sink strength of seeds. Consequently, it is worth investigating whether plants can be engineered to produce large seeds by modulating endosperm proliferation. Due to the importance of endosperm in seed development, considerable research effort has been devoted to elucidating the genetic control of endosperm development. This has enabled a biotechnological approach to be designed and tested in the model plant *Arabidopsis*, focussing on the over-expression of genes associated with increased endosperm proliferation using endosperm-specific promoters.

3.1.1 Genes associated with endosperm over-proliferation

The strategy outlined above requires genes that are responsible for endosperm over-proliferation in order to direct expression of their coding regions and potentially increase seed size and yield. Interploidy crosses between a diploid seed parent and a tetraploid pollen parent result in developing seeds with endosperm that has accelerated cell division, delayed cellularisation and an enlarged chalazal region due to the increased ratio of paternal to maternal genomes in the ('paternalised') seed (Scott *et al*., 1998). The gene expression underlying this parent-of-origin effect observed in paternalised seeds has been subjected to analysis by transcript profiling (Tiwari *et al*., unpublished). Genes found to be up-regulated in seeds with an excess of paternal genomes could be responsible for directing endosperm over-proliferation and therefore increased seed size. Using this list of genes, along with genes previously reported to be associated with increased cell proliferation in the literature, four candidate genes were selected: *PHERES 1* (*PHE1*, At1g65330), *AGAMOUS-LIKE 28* (*AGL28*, At1g01530), *AINTEGUMENTA* (*ANT*, At4g37750) and a gene encoding a putative sugar transporter (At1g77210).

Transcripts of the *PHE1* gene were found to be more abundant in paternalised seeds than those from a balanced cross between two diploid plants (Tiwari *et al*., unpublished). The *PHE1* coding region contains a highly conserved MADS box domain

which binds DNA. The *Arabidopsis* genome has 107 genes which possess a MADS box domain and represent a large family of transcription factors (Pařenicová *et al*., 2003). Although the function of many of these genes is unknown, some have been assigned roles in diverse aspects of development such as flowering time, floral meristem identity, floral organogenesis, fruit formation as well as seed pigmentation and endothelium development. *PHE1* is one of six genes known to be imprinted in *Arabidopsis* and is the only gene reported to be maternally repressed and paternally expressed (Kohler *et al*., 2005). Genomic imprinting is thought to have arisen as a consequence of conflict between the parents over resource allocation from the mother to the embryo (Haig and Westoby, 1989, 1991). The fitness of the mother is greatest when her resources are distributed equally amongst the offspring, while the father benefits when his offspring obtain a concentrated supply of maternal resources. Therefore, the model predicts that maternally and paternally derived alleles will have opposing effects on embryo growth, with growth promoters being paternally active and maternally silenced, and growth inhibitors being maternally active and paternally silenced. *PHE1* should therefore promote embryo and endosperm growth.

In wild-type plants, *PHE1* is expressed transiently from the paternal allele in both the embryo and endosperm following fertilisation (Kohler *et al*., 2003a). Repression of the maternal *PHE1* allele is mediated by the activity of the *FERTILISATION-INDEPENDENT SEED* (*FIS*) Polycomb-group (PcG) complex, which consists of four proteins encoded by *MEDEA* (*MEA*), *FERTILISATION-INDEPENDENT ENDOSPERM* (*FIE*), *FIS2* and *MULTICOPY SUPPRESSOR OF IRA 1* (*MSI1*) (Kohler *et al*., 2003b; Chanvivattana *et al*., 2004). The FIS complex regulates the methylation status of histones within the *PHE1* locus to determine gene expression (Makarevich *et al*., 2006; 2008). Loss-of-function of *MEA*, *FIE*, *FIS2* or *MSI1* genes leads to endosperm development in the absence of fertilisation and *PHE1* expression is elevated. Recently, UBIQUITIN-SPECIFIC PROTEASE 26 (UBP26) has been reported to be involved in the repression of the *PHE1* maternal allele, and may act in the same pathway as the FIS complex (Luo *et al*., 2008). The function of PHE1 is not fully understood as the downstream targets of this transcription factor are not yet determined and *phe1* mutants are indistinguishable from wild-type plants (Kohler *et al*., 2005). However, PHE1 has been shown to interact with AGL28, AGL40 and AGL62 suggesting that these four proteins are involved in the same developmental process (de Folter *et al*., 2005).

Like *PHE1*, *AGL28* was selected for this strategy as it was found to be up-regulated in paternalised seeds (Tiwari *et al*., unpublished). *AGL28* encodes a MADS box protein that interacts directly with PHE1 (de Folter *et al*., 2005). As well as a putative role in seed

development, *AGL28* has been shown to be involved in the regulation of flowering time (Yoo *et al*., 2006).

ANT encodes a transcription factor of the AP2-domain family that are only found in plants (Klucher *et al*., 1996; Elliot *et al*., 1996). *ANT* is highly expressed in ovule primordia and in the integuments, however *ANT* mRNA is also found in the primordia of other floral and vegetative organs (Elliot *et al*., 1996). The pattern of *ANT* expression is consistent with its function in ovule development, organ primordium initiation and organ growth. In *ant* mutants, the size of all lateral organs is reduced by a reduction in cell number. Conversely, in plants with ectopic *ANT* expression all lateral organs and developing embryos are enlarged due to increased cell numbers (Mizukami and Fischer, 2000). During organogenesis *ANT* mediates cell proliferation and organ growth by maintaining the meristematic competence of cells, prolonging the period in which cells are able to divide. Consequently, it is worth assessing whether over-expression of *ANT* in the endosperm would promote cell proliferation and increase seed size.

The final gene chosen encodes a putative sugar transporter (At1g77210, referred to here as *SUGAR*). The *SUGAR* gene was found to be up-regulated in paternalised seeds, with a 5-fold increase in the level of *SUGAR* transcripts compared to seeds from a balanced cross (Tiwari *et al*., unpublished) Therefore, *SUGAR* expression is positively correlated with the extent of endosperm proliferation. As the maternal and filial tissues form separate symplastic domains, the transfer of assimilate, primarily in the form of reducing sugars, into seeds requires membrane-localised transporter proteins (Stadler *et al*., 2005). The premise was that over-expressing the *SUGAR* gene in the endosperm would increase the amount of sugar transporters, thereby improving sugar uptake into the seed, increasing storage product accumulation and subsequently seed size and yield.

3.1.2 Endosperm-specific promoters

Several promoters active during the proliferative-phase of endosperm development in *Arabidopsis* have been identified and are essential for this approach (Tiwari *et al*., 2006), two of which have been utilised in the following experiments. The first promoter (referred to as *pPER*) is from the At5g46950 gene which encodes a putative invertase/pectin methylesterase inhibitor protein. *pPER* directs expression throughout the endosperm in peripheral, chalazal and micropylar regions, but not elsewhere in the plant. The second promoter (referred to as *pCZE*) is from the At5g07210 gene which encodes ARABIDOPSIS RESPONSE REGULATOR 21 (ARR21), one of eleven B-type two-component response regulators found in *Arabidopsis* (Tajima *et al*., 2004). *pCZE* targets gene expression solely to the chalazal region of the endosperm and is also active in pollen

(Tiwari *et al*., 2006). This chapter includes details of the cloning of two promoter::reporter constructs, *pPER::GUS* and *pCZE::GUS*, which were used to verify the activity of the endosperm-specific promoters.

The endosperm-specific *PER* promoter was employed to drive expression of the *PHE1*, *AGL28* and *ANT* coding regions as these genes are associated with endosperm over-proliferation and therefore should be targeted to the whole endosperm. In contrast, the *CZE* promoter was used to drive the *SUGAR* coding region, as the position of the chalazal endosperm at the maternal-filial interface suggests that this region has a more specialised role in nutrient transfer (Nguyen *et al*., 2000). Consequently, this chapter describes the cloning of four expression constructs: *pPER::PHE1*, *pPER::AGL28*, *pPER::ANT* and *pCZE::SUGAR*. These constructs were used to transform *Arabidopsis* and their effect on seed size and yield assessed.

3.2 Results

3.2.1 Promoter::reporter construct assembly and visualising promoter activity

Promoters active specifically within the endosperm were required to over-express genes associated with endosperm proliferation with the ultimate aim of increasing seed size and yield. To confirm the expression pattern of the *PER* and *CZE* promoters, histochemical localisation of ß-glucuronidase (GUS) activity was carried out. Previous analysis of the activity of these promoters used a genomic fragment from approximately 2 kb 5' of the presumed start codon to the end of the second intron (Tiwari *et al.* 2006). To fuse the promoters with alternative coding regions, it was necessary to remove the ATG and downstream gene sequence. The shortened *PER* and *CZE* promoters were amplified from genomic DNA (primers and PCR conditions are detailed in Table 2.3 and Table 2.4). Restriction site linkers were incorporated into the primers to enable directional cloning into the BJ60 vector. The promoter fragments were ligated into the pGEM-T vector and checked for sequence integrity. *Sal*I and *Kpn*I were used to excise the promoter fragments out of pGEM-T and transfer them to BJ60 upstream of the *uidA* (*GUS*) reporter gene. Finally, the promoter::*GUS* cassettes, along with the ocs 3' terminator, were then digested out of BJ60 using *Not*I and ligated into the binary vector, BJ40, completing the assembly of *pPER::GUS* and *pCZE::GUS* (Figure 3.1). The resulting reporter constructs were used to transform *Arabidopsis*.

Following transformation, GUS activity was assayed in seeds developing on T_1 generation, hemizygous plants. Three positive transformants, selected on kanamycin-containing media, were tested for each construct. Although levels of staining varied greatly, the expression patterns observed were identical to those published (Tiwari *et al*, 2006). The shortened *PER* promoter conferred GUS staining throughout the endosperm whereas the shortened *CZE* promoter gave staining that was restricted to the chalazal region (Figure 3.2).

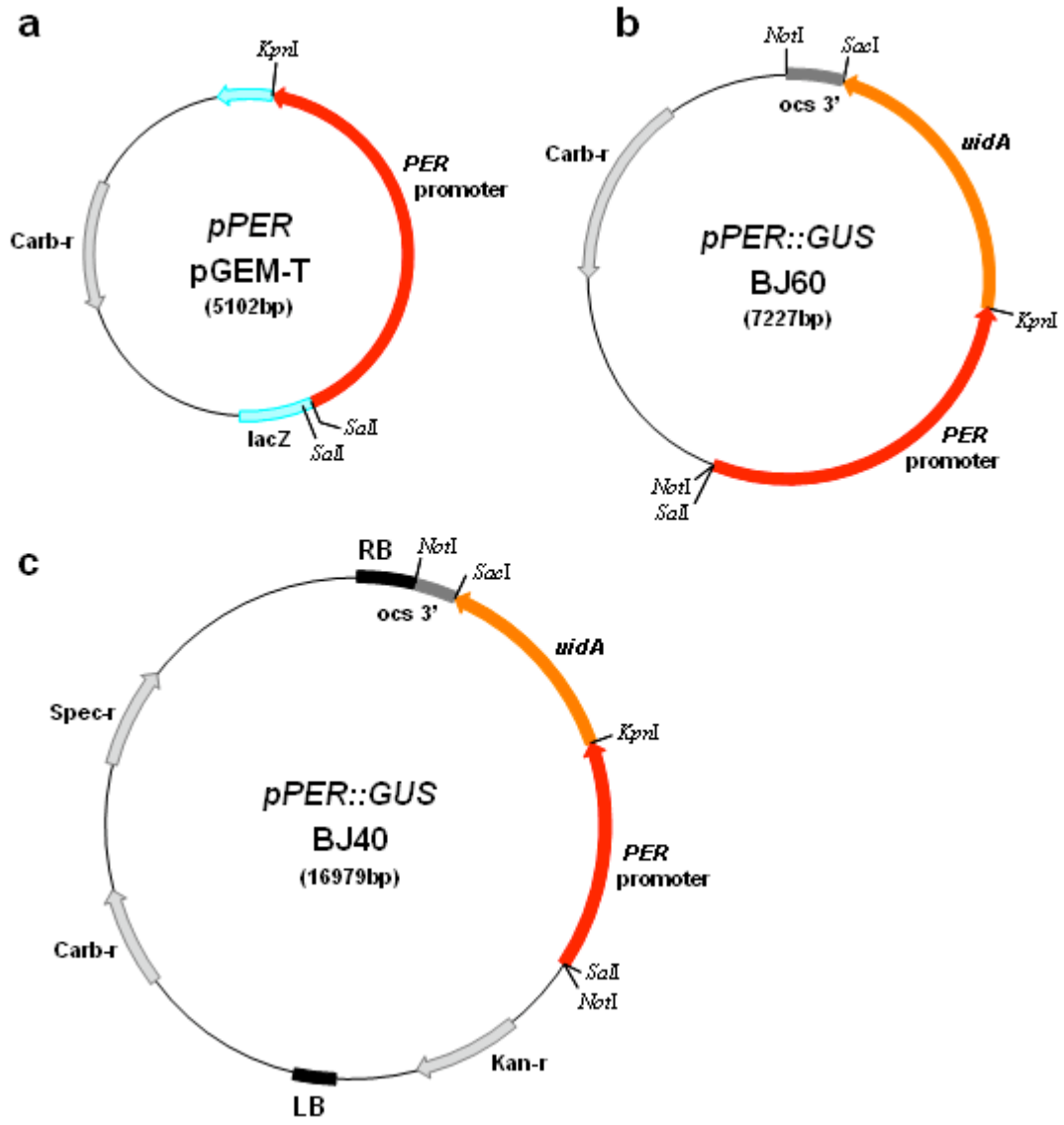

Figure 3.1: Assembly of the *pPER::GUS* reporter construct.

(a) The *PER* promoter was ligated into the pGEM-T vector. (b) The promoter was digested out of pGEM-T with *Sal*I and *Kpn*I and ligated into BJ60. (c) The promoter::GUS cassette, along with the ocs3' terminator, was lifted from BJ60 with a *Not*I digest and transferred to the binary vector, BJ40. (*pCZE::GUS* was created using the same method.)

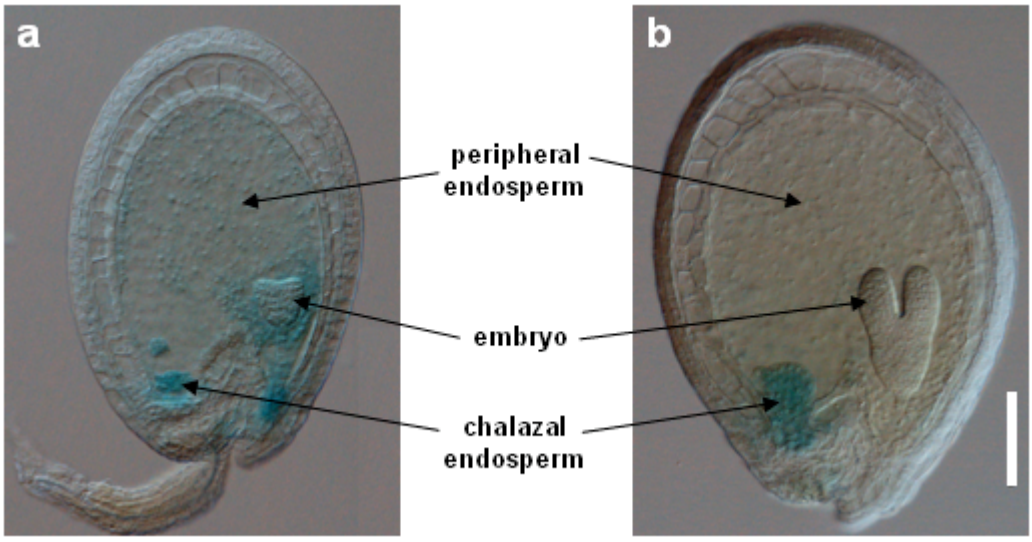

Figure 3.2: *GUS* expression directed by early endosperm-specific promoters.
Self-pollinated seeds of Col-0 plants transformed with the endosperm-specific reporter construct *pPER::GUS* (a), and the chalazal endosperm-specific reporter construct *pCZE::GUS* (b). Bar = 100 μm.

3.2.2 Assembly of the endosperm-specific over-expression vectors and verification of transgenic Arabidopsis

The coding regions of *PHE1*, *AGL28* and *ANT* were amplified by PCR from a cDNA template generated from the RNA of 3 DAP siliques (primers and PCR conditions are detailed in Table 2.3 and Table 2.4). The *SUGAR* coding sequence was amplified from 5 DAP siliques. Restriction site linkers were incorporated into the primers to enable directional cloning into the shuttle vector BJ36. Prior to insertion into BJ36, each coding region was cloned into pGEM-T and checked for sequence integrity. The endosperm-specific promoters (Section 3.2.1) were transferred first into BJ36 using the *Sal*I and *Kpn*I restriction sites. *Kpn*I and *Xma*I were then used to lift the coding regions from pGEM-T and position them in the correct orientation downstream of the appropriate promoter. Subsequently each expression cassette, consisting of promoter, coding region and ocs 3' terminator, was cut out of BJ36 with *Not*I and inserted into BJ40. For an example of the vector construction process, see Figure 3.3. The four completed constructs were referred to as *pPER::PHE1*, *pPER::AGL28*, *pPER::ANT* and *pCZE::SUGAR*.

The endosperm-specific over-expression vectors also introduced a kanamycin resistance gene into *Arabidopsis*, enabling selection of positive transformants using kanamycin-containing media. Positive seedlings were screened by PCR to confirm the presence of the transgene (primers and PCR conditions are detailed in Table 2.1 and Table 2.2). Homozygous, single insertion lines were selected for detailed phenotype analysis.

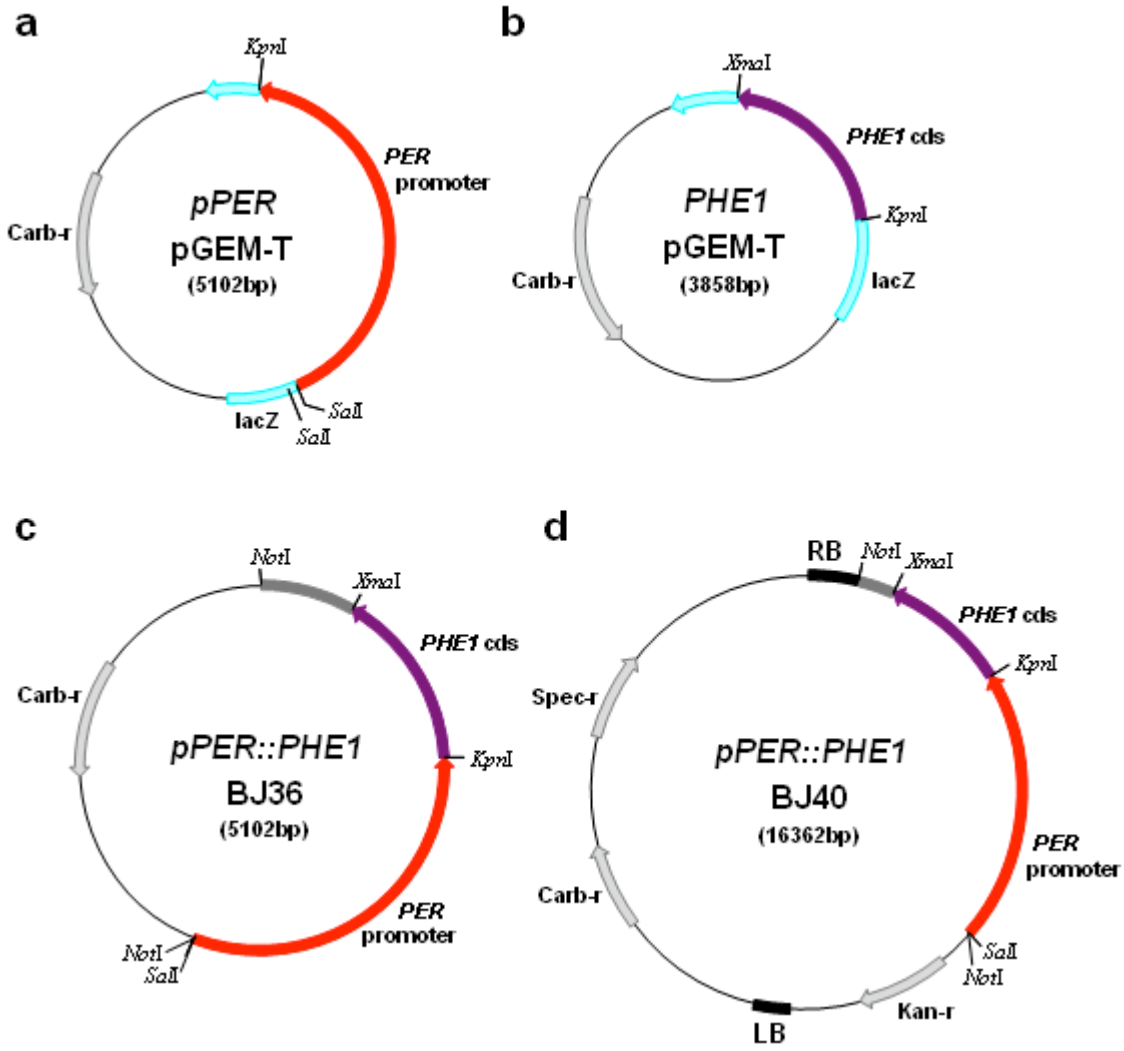

Figure 3.3: Assembly of the *pPER::PHE1* expression vector.

The *PER* promoter (a) and the *PHE1* coding sequence (b) were independently ligated into the pGEM-T vector. (c) The endosperm-specific promoter was digested out of pGEM-T with *Sal*I and *Kpn*I and ligated into BJ36. Following this, the *PHE1* cds was lifted out of pGEM-T with *Kpn*1 and *Xma*I and cloned into BJ36 alongside the promoter. (d) The *pPER::PHE1* cassette, along with the ocs3' terminator, was lifted from BJ36 with a *Not*I digest and transferred to the binary vector, BJ40. (The expression constructs *pPER::AGL28*, *pPER::ANT* and *pCZE::SUGAR* were generated using the same method.)

3.2.3 Seed size analyses of pPER::PHE1 plants

To determine whether over-expression of *PHE1* in the endosperm altered seed size, three independent *pPER::PHE1* transformation lines that were homozygous for a single insertion were analysed for both embryo sac area and seed weight. The embryo sac area refers to the cavity surrounded by the seed coat which contains the chalazal endosperm cyst and the embryo. Only seeds orientated so that the embryo and the chalazal endosperm cyst were equally visible were used. The embryo sac area of developing seeds with heart and torpedo stage embryos was measured in order to provide an additional indicator of increased seed size. Unlike seed weight, which represents the size of mature seeds, the embryo sac area represents the size of developing seeds and is an estimate of the space occupied by the developing endosperm.

No significant difference was found between the embryo sac area of *pPER::PHE1* and wild-type seeds containing heart stage embryos (ANOVA followed by Tukey's multiple comparisons, $P = 0.398$) (Figure 3.4a). In contrast, the embryo sac area of torpedo stage seeds from the *pPER::PHE1*-10.4 line was 15.6% larger than that of the wild-type, and this difference was statistically significant (Student's *t*-test; Col-0 vs. 10.4, $P < 0.001$). However, *pPER::PHE1*-9.3 and *pPER::PHE1*-16.1 did not show a significant difference.

Two methods were used to compare seed weight in *pPER::PHE1* plants with those of wild-type. In the first treatment, only six flowers on the primary inflorescence of each plant were allowed to set seed (restricted pollination); this forced all plants to produce approximately the same number of seeds, allowing a controlled, although artificial, comparison of seed size. As the seeds from each silique were collected separately, restricted pollinations also enabled the assessment of seed number/silique. In the second treatment, all flowers were allowed to set seed (unrestricted pollination).

In unrestricted pollinations, no difference was detected between the seed weight of *pPER::PHE1* lines and wild-type (ANOVA followed by Tukey's multiple comparisons, $P = 0.146$) (Figure 3.4b). In restricted pollinations, none of the transgenic lines produced seeds that were significantly different in weight from wild-type (Figure 3.4c). The fertility of plants containing the *pPER::PHE1* transgene was not compromised as more than 60 seeds were produced per silique and there was no significant difference between the seed set of transgenic lines and wild-type (ANOVA followed by Tukey's multiple comparisons, $P = 0.274$) (Figure 3.4d). Full fertility is important as any seed size increase would subsequently have a potentially positive effect on total seed yield.

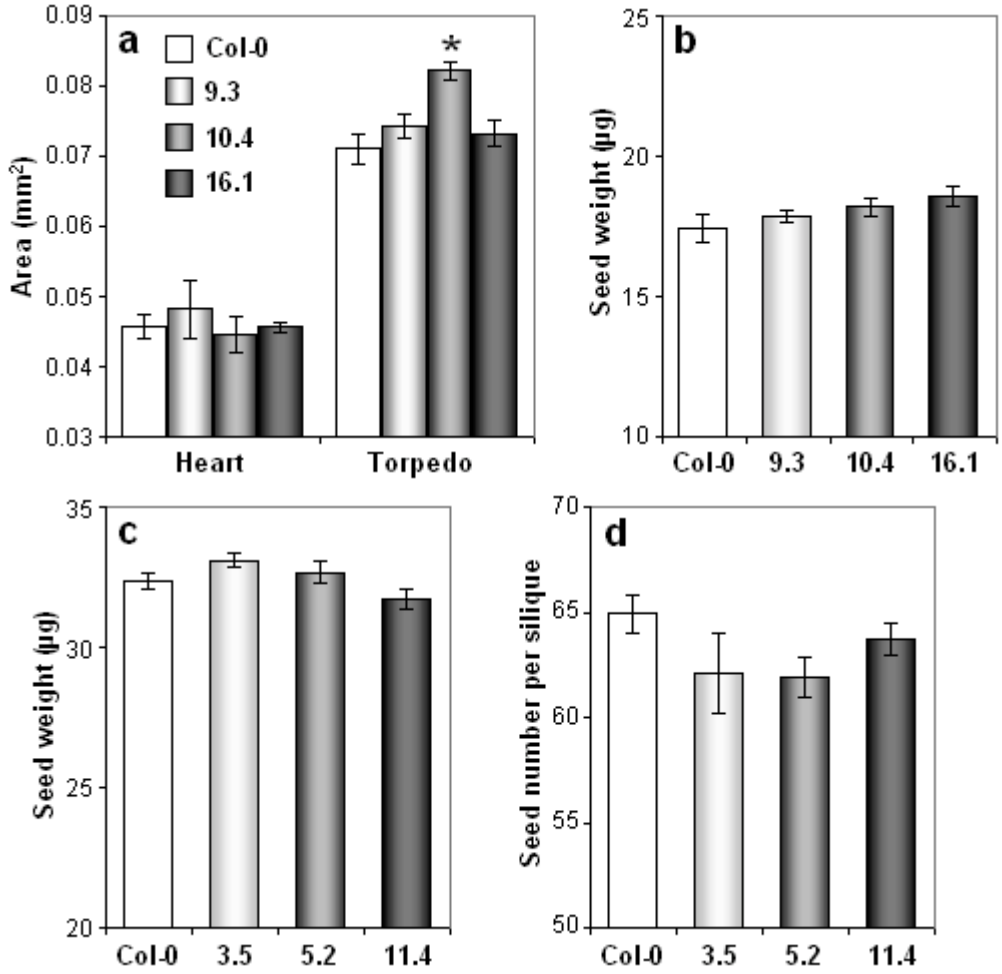

Figure 3.4: Seed phenotype analysis of three independent, homozygous, single insertion *pPER::PHE1* lines compared to Col-0.

(a) Comparison of embryo sac area at two different embryonic growth stages ($n \geq 14$ seeds). Mean seed weight in unrestricted (b) and restricted (c) pollinations ($n \geq 5$ plants). (d) Seed set as recorded in restricted pollinations ($n \geq 7$ plants). For each set of comparisons, values that differ at the 0.05 significance level from wild-type are labelled with a *. Error bars = s.e.m.

3.2.4 Seed size analyses of pPER::AGL28 plants

To determine whether over-expression of *AGL28* in the endosperm altered seed size, four independent *pPER::AGL28* transformation lines that were homozygous for a single insertion were analysed for increased embryo sac area, chalazal endosperm area and seed weight. The embryo sac area of plants containing the *pPER::AGL28* transgene was significantly larger than wild-type at 6 DAP in all four lines analysed (Figure 3.5a), and represented a 7.8-18.5% increase (Student's *t*-test; Col-0 vs. 4.4, $P = 0.027$; Col-0 vs. 8.2, $P < 0.001$; Col-0 vs. 11.1, $P < 0.001$; Col-0 vs. 17.5, $P = 0.002$). No statistical difference was detected between the chalazal endosperm area of the *pPER::AGL28* lines and Col-0 seeds at 6 DAP (ANOVA followed by Tukey's multiple comparisons, $P = 0.390$) (Figure 3.5b).

Although the embryo sac area was found to be larger in *pPER::AGL28* plants than in wild-type this did not translate into a higher final seed weight. The mean seed weight of plants containing the *pPER::AGL28* transgene was not significantly different from wild-type in any of the lines tested, in either unrestricted (ANOVA followed by Tukey's multiple comparisons, $P = 0.139$) or restricted (ANOVA followed by Tukey's multiple comparisons, $P = 0.204$) pollinations (Figure 3.5c,d). Although no statistical differences were found, the *pPER::AGL28*-11.1 line produced the heaviest seeds in both unrestricted and restricted pollinations and also the largest embryo sac area. The *pPER::AGL28* transgene had no effect on fertility as measured by the number of seeds per silique (Mann-Whitney *U*-test, Col vs. 4.4, $P = 0.052$) (Student's *t*-test, Col vs. 8.2, $P = 0.358$; Col vs. 11.1, $P = 0.167$; Col vs. 17.5, $P = 0.229$) (Figure 3.5e).

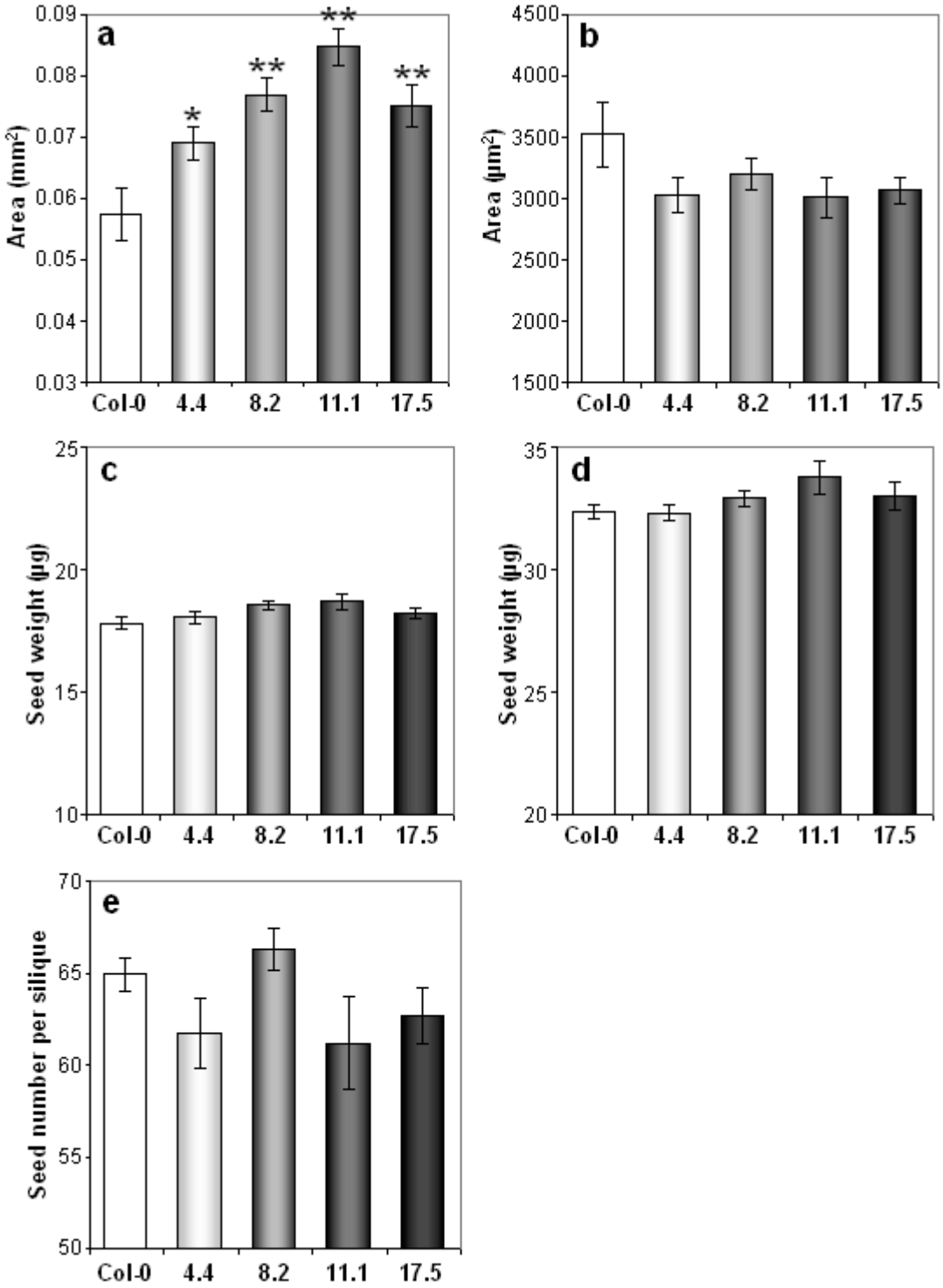

Figure 3.5: Seed phenotype analysis of four independent, homozygous, single insertion *pPER::AGL28* lines compared to Col-0.

(a) Comparison of embryo sac area at 6 DAP (n ≥ 28 seeds). (b) Comparison of chalazal endosperm area at 6 DAP (n ≥ 9 seeds). Mean seed weight in unrestricted (c) and restricted (d) pollinations (n ≥ 5 plants). (e) Seed set as recorded in restricted pollinations (n ≥ 7 plants). For each set of comparisons, values that differ from wild-type at the 0.05 significance level are labelled with * and at the 0.01 significance level with **. Error bars = s.e.m. Area measurements and unrestricted seed weight analysis were performed by Zhou Ye (University of Bath).

3.2.5 Seed size analyses of pPER::ANT plants

Three independent, homozygous, single insertion *pPER::ANT* lines were analysed for seed weight in unrestricted pollinations (Figure 3.6a). All lines produced seeds that were significantly different in weight to wild-type (Mann-Whitney *U*-test, Col vs. 4.1, P = 0.001) (Student's *t*-test, Col vs. 5.7, P < 0.001; Col vs. 15.5, P = 0.028; Col vs. 16.4, P < 0.001). However, *pPER::ANT*-4.1 and *pPER::ANT*-15.5 produced heavier seeds whereas and *pPER::ANT*-16.4 produced lighter seeds than wild-type. Work was discontinued on *pPER::ANT*-4.1 as siliques failed to fully elongate. As this phenotype was only present in a single transformation line it was deemed to be caused by the insertion of the transgene rather than a direct effect of transgene expression. Consequently, only two *pPER::ANT* lines were assessed for seed weight in restricted pollinations (Figure 3.6b, Figure 3.7). Both *pPER::ANT*-15.5 and *pPER::ANT*-16.4 produced heavier seeds, 15.4% and 3.7% respectively, than the wild-type and these differences were significant (Student's *t*-test, Col vs. 15.5, P < 0.001; Col vs. 16.4, P = 0.018). Fertility was not affected by introduction of the *pPER::ANT* transgene as the number of seeds per silique was not significantly different from wild-type (ANOVA followed by Tukey's multiple comparisons, P = 0.351) (Figure 3.6c). Therefore, the increase in seed weight over wild-type observed in restricted pollinations was not a result of reduced seed number.

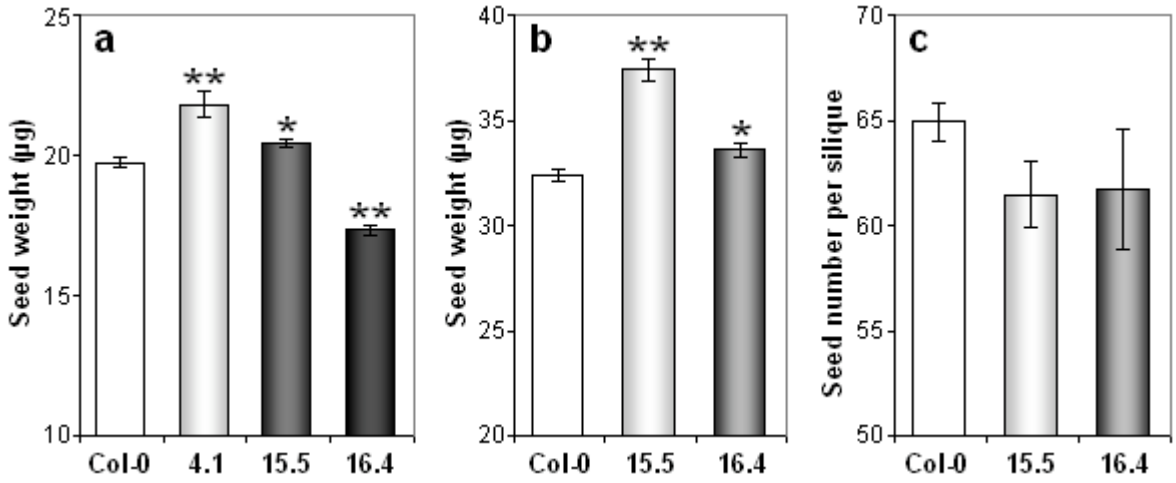

Figure 3.6: Seed weight analysis of independent, homozygous, single insertion *pPER::ANT* lines compared to Col-0.

Mean seed weight in unrestricted (a) and restricted (b) pollinations (n ≥ 6 plants). (c) Seed set as recorded in restricted pollinations (n ≥ 6 plants). For each set of comparisons, values that differ from wild-type at the 0.05 significance level are labelled with * and at the 0.01 significance level with **. Error bars = s.e.m. Unrestricted seed weight was determined by Zhou Ye (University of Bath).

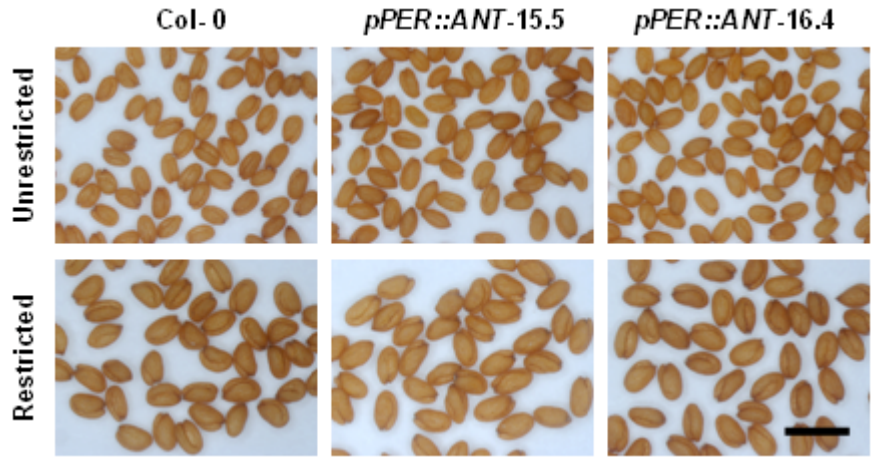

Figure 3.7: Seed phenotype of *pPER::ANT* plants

Photographs of Col-0, *pPER::ANT* 15.5 and *pPER::ANT* 16.4 seeds produced in unrestricted and restricted pollinations. Bar = 1 mm.

3.2.6 Seed size analyses of pCZE::SUGAR plants

In order to test whether introduction of the *pCZE::SUGAR* construct into *Arabidopsis* increased endosperm growth and seed size, four independent, homozygous, single insertion lines were analysed for mean seed weight and seed set (Figure 3.8). In unrestricted pollinations all four transgenic lines produced seeds that were smaller than wild-type, and this difference was significant (Mann-Whitney *U*-test, Col vs. 2.5, $P = 0.001$) (Student's *t*-test, Col vs. 5.5, $P < 0.001$; Col vs. 7.3, $P < 0.001$; Col vs. 11.7, $P < 0.001$). In contrast, in restricted pollinations no difference was observed between the seed size of wild-type and any of the *pCZE::SUGAR* lines (ANOVA followed by Tukey's multiple comparisons, $P = 0.795$). Fertility was not affected by introduction of the *pCZE::SUGAR* transgene in three of the four lines tested, however *pCZE::SUGAR*-5.5 had significantly fewer seeds per silique compared to wild-type (Student's *t*-test, Col vs. 5.5, $P < 0.001$).

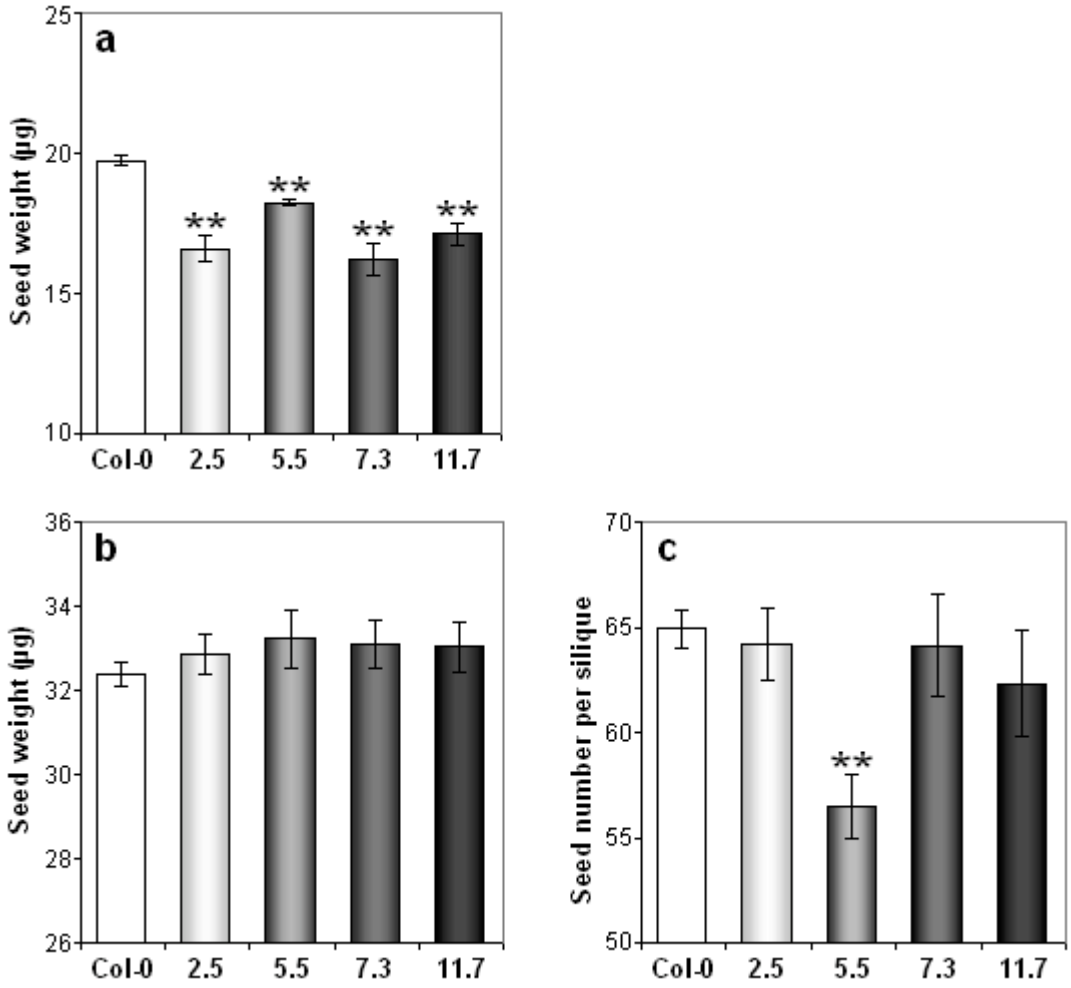

Figure 3.8: Seed weight analysis of four independent, homozygous, single insertion *pCZE::SUGAR* lines compared to Col-0.

Mean seed weight in unrestricted (a) and restricted (b) pollinations (n ≥ 6 plants). (c) Seed set as recorded in restricted pollinations (n ≥ 6 plants). For each set of comparisons, values that differ from wild-type at the 0.01 significance level are labelled with **. Error bars = s.e.m. Unrestricted seed weight was determined by Zhou Ye (University of Bath).

3.2.7 Expression of the sugar transporter in pCZE::SUGAR plants

Semi-quantitative RT-PCR was carried out to determine whether the *pCZE::SUGAR* transgene was being expressed and also if this expression enhanced the endogenous level of *SUGAR* mRNA present in the endosperm. This would ensure that any phenotype observed was a result of *SUGAR* over-expression. RNA was extracted from 5 DAP siliques in order to generate cDNA. The level of transgene expression was determined using a primer 'SUGAR RT-PCR For' which anneals within the gene sequence and another 'pFGC5941 ocs R2' that anneals within the ocs 3' terminator (Figure 3.9a) (primers and PCR conditions are detailed in Table 2.1 and Table 2.2). The overall *SUGAR* expression level (endogenous plus transgene-driven expression) was determined using two primers

that anneal within the gene coding sequence, 'SUGAR RT-PCR For2' and 'SUGAR RT-PCR Rev' (Figure 3.9b). The forward primer bridges an exon-exon boundary to prevent amplification of any genomic DNA contaminants. Three independent transformation lines were analysed. All tested positive for expression of the *pCZE::SUGAR* transgene, although a range of expression levels was found (Figure 3.9c). Furthermore, the endogenous level of *SUGAR* mRNA was considerably increased in all three lines compared to wild-type.

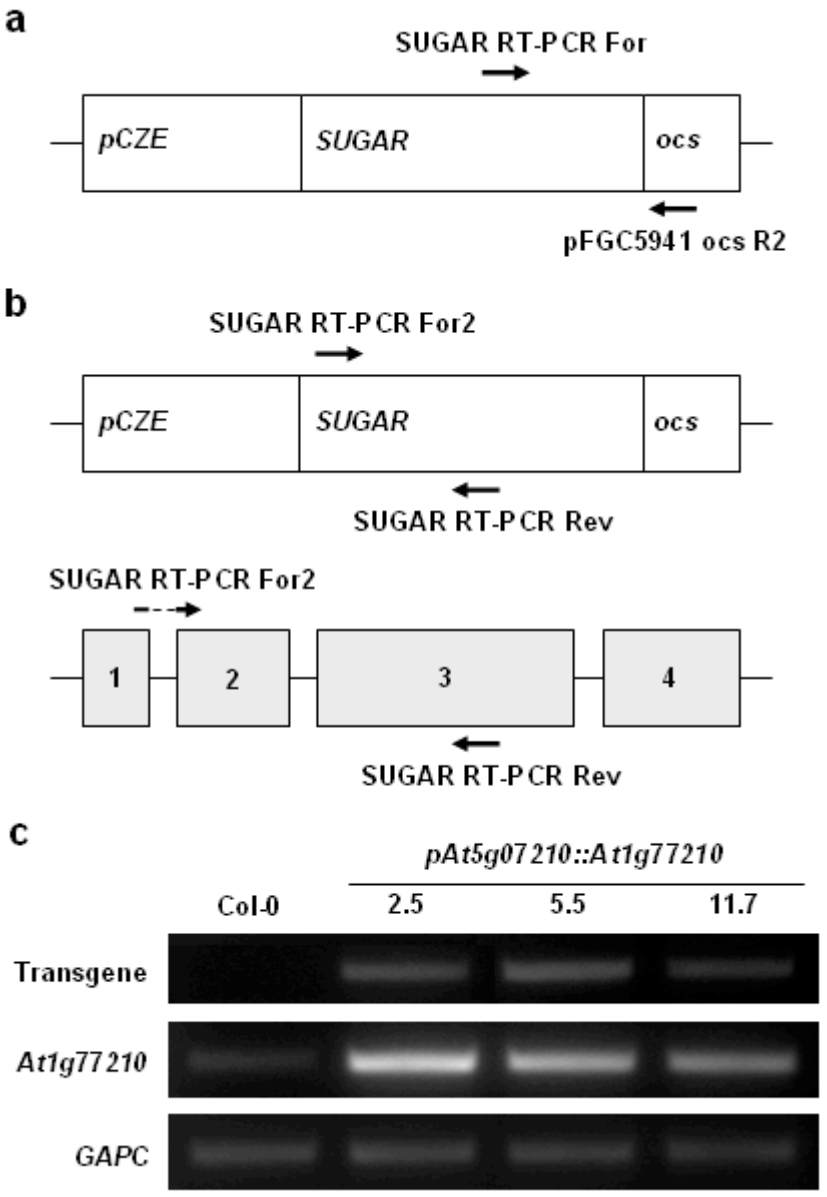

Figure 3.9: The *pCZE::SUGAR* transgene is expressed in plants and enhances the endogenous levels of *SUGAR* mRNA.

(a) Primer orientation for amplification of the *SUGAR* transgene. (b) Primer orientation for amplification of total *SUGAR* transcripts. Grey boxes represent exons within the *ARF2* coding region, with exon numbers given. (c) Semi-quantitative RT-PCR on RNA extracted from 5 DAP siliques of Col-0 and three independent, homozygous, single insertion *pCZE::SUGAR* lines.

3.3 Discussion

This chapter describes a biotechnological approach to engineer large seeds through the modulation of endosperm development. Genes associated with increased cell proliferation were over-expressed specifically within the endosperm of *Arabidopsis* Col-0 plants. The phenotype of the transgenics was assessed for increased seed size and several positive outcomes were observed.

3.3.1 Over-expression of transcription factors associated with endosperm proliferation has the potential to increase seed size

The *pPER::PHE1*, *pPER::AGL28* and *pPER::ANT* constructs were introduced into *Arabidopsis* with the aim of enhancing endosperm growth and subsequently, seed size. These constructs were designed to over-express genes associated with increased cell proliferation throughout the endosperm using the *PER* promoter. It is not a coincidence that the three genes selected for this approach encode transcription factors. PHE1, AGL28 and ANT proteins can bind DNA and alter the expression of other genes. It is possible that they act near or at the top of a transcriptional cascade which promotes seed growth; therefore by increasing the activity of a single gene, a multitude of cellular changes could be induced that would result in the production of large seeds.

In plants ectopically-expressing the *ANT* transcription factor, lateral organs and developing embryos are enlarged due to increased cell numbers (Mizukami and Fischer, 2000). During organogenesis ANT mediates cell proliferation and organ growth by maintaining the meristematic competence of cells, therefore endosperm-specific over-expression of *ANT* had the potential to increase endosperm proliferation and seed size. Variation was found between different transgenic lines containing the *pPER::ANT* construct, but importantly *pPER::ANT*-15.5 showed significantly increased seed size compared to wild-type in both restricted and unrestricted pollinations (Figure 3.6). The fertility of *pPER::ANT* plants was not compromised suggesting that this increase in seed size is not associated with a reduction in seed number but is due to altered endosperm development. *pPER::ANT*-15.5 seeds were 15.4% larger in restricted pollinations and 3.4% larger in unrestricted pollinations than wild-type seeds. The greater percentage increase over wild-type seed size in restricted pollinations is not surprising since more resources are available per seed. This data suggests that introduction of the *pPER::ANT* transgene into plants has increased the sink strength of developing seeds, and as a result the total seed yield of these transgenic plants may be improved.

Transcriptional profiling revealed that transcripts encoding the PHE1 transcription factor were more abundant in large seeds with over-proliferated endosperm (Tiwari *et al*., unpublished). However, plants transformed with the *pPER::PHE1* construct showed no consistent alterations to seed size. Neither the embryo sac area nor the seed weight in restricted and unrestricted pollinations was increased (Figure 3.4). Possible explanations for this include: introduction of the *pPER::PHE1* transgene did not lead to a substantial increase in *PHE1* expression; accumulation of *PHE1* mRNA did not lead to an equivalent accumulation of PHE1 protein due to post-transcriptional gene control; a high level of PHE1 protein may not be sufficient to alter endosperm proliferation; increased endosperm proliferation failed to influence final seed size. Further work should be carried out to identify which of these explanations is responsible.

PHE1 is a maternally repressed, paternally expressed imprinted gene (Kohler *et al*., 2005); therefore according to the parental conflict model (Haig and Westoby, 1989, 1991) *PHE1* should encode a growth promoter. The *pPER::PHE1* transgene introduces an additional copy of the *PHE1* coding region into the *Arabidopsis* genome under the control of the endosperm-specific *PER* promoter. Establishment of *PHE1* imprinting requires binding of the FIS complex to a Polycomb response element (PRE) in the promoter region of *PHE1* and a differentially methylated region (DMR) situated downstream of the *PHE1* locus to be methylated (Makarevich *et al*., 2008). The extra copy should not be affected by genomic imprinting as it does not possess the necessary regulatory elements hence plants containing *pPER::PHE1* should have elevated *PHE1* expression in the endosperm. If *PHE1* is a growth promoter, up-regulation of this gene should in theory promote embryo growth and increase seed size.

Like *PHE1*, transcripts of the *AGL28* gene were also over-represented in large, paternalised seeds (Tiwari *et al*., unpublished). The embryo sac area of 6 DAP developing seeds of *pPER::AGL28* plants was enlarged; however this was not associated with an increase in final seed size, as in both restricted and unrestricted pollinations there was no significant difference in seed weight compared to wild-type (Figure 3.5). This result is surprising, since developing seeds reach their maximal length at approximately 6-7 DAP (Alonso-Blanco *et al*., 1999), and suggests that seeds containing the *pPER::AGL28* transgene expand more rapidly than wild-type but once achieving their maximum size cease to grow further.

The failure to increase final seed size by transforming plants with either *pPER::PHE1* or *pPER::AGL28* may have arisen for many reasons. One possible explanation takes into account that PHE1 and AGL28 interact directly with AGL40 and AGL62, suggesting that these four proteins are involved in the same developmental

process and may operate together in a complex (de Folter *et al*., 2005). Although little is known about the function of AGL28 and AGL40 in seed development, AGL62 has been the subject of a recent study. *AGL62* is expressed exclusively in the endosperm during seed development and its expression declines abruptly before the syncytial endosperm becomes cellularised (Kang *et al*., 2008). Mutant *agl62* seeds cellularise prematurely, have fewer endosperm nuclei and contain abnormal embryos. The *AGL62* expression pattern and loss-of-function mutant phenotype indicate that AGL62 is required for suppression of cellularisation during the syncytial phase of endosperm development. This study also revealed that the FIS complex may mediate *AGL62* activity, although in contrast to *PHE1* this gene is not imprinted. The timing of endosperm cellularisation is important as it influences the extent of endosperm proliferation, sink strength and seed size (Scott *et al*., 1998; Garcia *et al*., 2003). This evidence suggests that PHE1 and AGL62 at least are involved in similar processes and further research in this area may reveal roles for both AGL28 and AGL40 in seed development. If these four proteins function as a complex, over-expression of only a single component may not be sufficient to increase seed size, as all four proteins need to be more abundant to create more functional complexes.

3.3.2 The role of a sugar transporter in seed size determination

As *SUGAR* was found to be up-regulated in the large seeds produced in interploidy crosses between a diploid seed parent and a tetraploid pollen parent (Tiwari *et al*., unpublished), it was worth testing whether over-expression of this putative sugar transporter could increase sugar uptake and final seed size. The results presented here show that in unrestricted pollinations seeds produced by *pCZE::SUGAR* plants have a significantly lower seed weight than wild-type; whilst in restricted pollinations no statistical difference in seed weight was found (Figure 3.8). Semi-quantitative RT-PCR revealed that plants possessing the *pCZE::SUGAR* transgene have elevated levels of *SUGAR* mRNA in 5 DAP siliques (Figure 3.9). Although over-expression of the *SUGAR* gene has been achieved, it is not certain whether the increase in *SUGAR* mRNA is associated with an increase in functional *SUGAR* transporter protein. If introduction of the *pCZE::SUGAR* transgene did lead to the accumulation of more membrane-localised sugar transporters in the chalazal endosperm, it is possible an increase in sugar uptake may not be possible due to limited availability of sugars from the mother plant.

The reduction in seed size observed here in unrestricted pollinations of *pCZE::SUGAR* plants is surprising, however it could be explained by modifications in sugar signalling that have been caused by increased *SUGAR* expression. The *SUGAR* gene encodes a putative sugar transporter but it is not known which particular sugars it carries

across cell membranes. In most plants, sucrose is the major transported carbon source and additionally it is the main sugar transported into seeds, prior to its conversion to fructose and glucose by invertases (Weber *et al*., 1998). Sugars may be metabolised into storage products but can also act as signalling molecules (Wobus and Weber, 1999; Sheen *et al*., 1999). Sucrose may play a role in positively regulating the cell cycle and thereby determine the extent of cell division (Riou-Khamlichi *et al*., 2000). During seed development, the ratio of hexoses to sucrose controls the transition from rapid cell division to storage product accumulation (Weber *et al*., 1998). It is evident that sugars play many important roles in developing seed which may potentially be perturbed by over-expression of the *SUGAR* transporter.

In *Arabidopsis*, the SUC5 sucrose transporter is the only sucrose transporter reported to be active specifically during seed development (Baud *et al*., 2005). The *SUC5* expression pattern begins at the micropylar pole at 3-4 DAP and is later also found at the chalazal pole. Developing *suc5* mutant seeds show a strong but transient reduction in FA concentration, altered FA profiles and also delayed embryo development at 8 DAP. This highlights the importance of the SUC5 transporter in nutrient uptake but also indicates that other mechanisms of transporting sugar are involved as *suc5* mutants produce viable seeds in which endosperm development is not compromised.

Since reduced activity of a sucrose carrier can lower nutrient uptake into seeds, it is plausible that introduction of the *pCZE::SUGAR* cassette into *Arabidopsis* Col-0 could increase the activity of the SUGAR transporter, enhance the unloading of assimilate into seeds, and increase seed size. Sucrose may pass from the phloem to the embryo via many possible routes. The organisation of the chalazal endosperm suggests that it is specialised for the uptake and processing of metabolites before releasing them into the developing seed (Nguyen *et al*., 2000). However, a recent study in *B. napus* suggests that sucrose is more likely to take a shorter route, either directly through the suspensor or the micropylar endosperm to the embryo (Morley-Smith *et al*., 2008). Furthermore, the SUC5 transporter is confined to the mycropylar endosperm until the embryo has reached torpedo-stage (Baud *et al*., 2005). It is possible that over-expression of *SUGAR* within the chalazal endosperm increases sugar uptake from the maternal tissues into the endosperm but transfer from the endosperm to the embryo is limiting. Further research is required to characterise this particular sugar transporter and its role in seed development.

3.3.3 Further Work

Further analyses of the transgenic plants are necessary to confirm that any phenotypes observed are due to the introduction of the transgene. Semi-quantitative RT-PCR

confirmed that introduction of the *pCZE::SUGAR* transgene caused an increase in the level of *SUGAR* mRNA. The over-expression of *PHE1*, *AGL28* and *ANT* should be verified using the same method. The *pPER::PHE1*, *pPER::AGL28* and *pPER::ANT* constructs were introduced into *Arabidopsis* in order to enhance seed size through increasing endosperm proliferation. Therefore, the extent of endosperm proliferation, i.e. the number of endosperm nuclei, and also the timing of endosperm cellularisation should be determined in the transgenic plants and compared to wild-type seeds. Large seeds were produced by plants containing the *pPER::ANT* transgene, however more independent transgenic lines should be tested to validate these results. A yield assessment of *pPER::ANT* plants may now be carried out as the transgene had no effect on fertility.

Individually, the introduction of the *pPER::PHE1* and *pPER::AGL28* constructs into plants did not affect final seed weight. As the *PHE1*, *AGL28*, *AGL40* and *AGL62* gene products may function as a protein complex, plants over-expressing all four of these genes should be generated and assessed for seed size and yield in order to test this hypothesis.

3.3.4 Summary

A transgenic approach was used to engineer *Arabidopsis* with large seeds and thereby increased seed yield. The seed phenotype of plants containing the *pPER::PHE1* transgene remained similar to wild-type. Seeds of *pPER::AGL28* had a larger embryo sac area, although final seed size was not increased. *pPER::ANT* plants produced seeds that were significantly larger than wild-type in both restricted and unrestricted pollinations. Surprisingly, an increase in *SUGAR* transcripts through introduction of the *pCZE::SUGAR* construct led to a decrease in seed size in unrestricted pollinations; however in restricted pollinations seed weight was not different from wild-type. Importantly, altering the expression of these four genes specifically within the endosperm did not affect fertility.

Table 3.1: Results summary for transgenic plants containing endosperm-specific over-expression constructs.

Over-expression construct	Embryo sac area	Chalazal endosperm area	Mean seed weight		Fertility	Gene expression
			Restricted	Unrestricted		
pPER::PHE1	#	ND	#	#	#	ND
pPER::AGL28	>	#	#	#	#	ND
pPER::ANT	ND	ND	>	>	#	ND
pCZE::SUGAR	ND	ND	#	<	#	>

ND = not determined. > = significant increase compared to wild-type. < = significant decrease compared to wild-type. # = not significantly different from wild-type.

4. ENGINEERING ENLARGED SEED SIZE USING INTEGUMENT-LED APPROACHES

4.1 Introduction

Development of the ovule integuments, and subsequently the seed coat, can significantly influence seed size (Egli, 1990; Weber *et al*., 1996; Garcia *et al*., 2005; Schruff *et al*., 2006), therefore they present an attractive target for biotechnological approaches to increase seed size and yield. Recently, the *auxin response factor 2* (*arf2*)/ *megaintegumenta* (*mnt*) mutant with altered integument development was identified in the Col-0 ecotype of *Arabidopsis* which provides an opportunity to test whether seed coat modification can enhance seed size and yield. Ultimately, if successful this would provide technology to raise yields in economically valuable crop species such as oilseed rape or even rice, maize and wheat.

4.1.1 arf2 mutants have enlarged seeds

The *Arabidopsis arf2* mutant produces large, pointed seeds associated with extra cell divisions in the ovule integuments prior to fertilisation (Schruff *et al*., 2006). As the integuments are entirely derived from maternal tissues it is not unexpected that the *arf2* mutation has a maternal effect on seed size; hence regardless of whether a homozygous mutant mother plant is fertilised by *arf2* or a wild-type plant, large seeds are always produced. Aside from altered seed development, the *arf2* mutant displays a pleiotropic vegetative and floral phenotype (Okushima *et al*., 2005; Schruff *et al*., 2006). Extra cell division in many organs results in a mutant with large leaves, thick and twisted stems, and reduced seed set due to enlarged sepals and gynoecia which prevent flower opening and self-pollination (Schruff *et al*., 2006). Stamens are short and pollen is deposited on the side of the over-elongated gynoecium rather than on the stigma, probably as the failure of flower opening restricts elongation of the stamen filaments. In addition, *arf2* mutants are delayed in the onset of flowering, leaf senescence, floral organ abscission and silique ripening (Ellis *et al*., 2005).

ARF2 is a member of a large family of 23 transcription factors in *Arabidopsis* that regulate gene expression in response to the plant hormone, auxin (Guilfoyle and Hagen, 2007). Several other *ARF* genes have also been characterised: *ARF3*/*ETTIN* is required for correct patterning of the gynoecium (Sessions and Zambryski, 1995; Sessions *et al*., 1997); *ARF5*/*MONOPTEROUS* is involved in vascular differentiation and embryo patterning (Berleth and Jürgens, 1993; Hardtke and Berleth, 1998); both *ARF6* and *ARF8* regulate

flower maturation (Nagpal *et al*., 2005); and *ARF7*/*NON-PHOTOTROPHIC HYPOCOTYL4* operates in differential growth responses in the hypocotyls (Stowe-Evans *et al*., 1998; Harper *et al*., 2000). Since the *arf2* mutant exhibits increased organ size due to cell over-proliferation, it is believed that the wild-type ARF2 protein functions as a general repressor of cell division (Schruff *et al*., 2006).

Many different *arf2* mutant alleles have been isolated. The EMS-induced *mnt*/*arf2-9* allele isolated in the Scott lab (Schruff *et al*., 2006) and utilised in the experiments described in this chapter, was originally discovered in a seed size screen (Adams, 2002). A single base change from G to A at position 665 from the translational start results in altered splicing and the deletion of four bases leading to a frameshift at amino acid position 123 and a premature stop codon at position 167. These alterations occur in the DNA binding domain of ARF2 and probably cause a complete loss of function (Schruff *et al*., 2006).

4.1.2 Reduced fertility in arf2 mutants artificially increases seed size

Due to the failure of flowers to open and self-pollinate, seed set is dramatically reduced in *arf2* mutants (Schruff *et al*., 2006). Low seed number artificially raises seed weight as more resources are available per seed (Alonso-Blanco *et al.,* 1999; Borrás and Otegui, 2001). However, *arf2* seeds were shown to be up to 21% heavier than wild-type when seed set was held constant (Schruff *et al*., 2006). Therefore, it is likely that increased seed size in *arf2* is in part due to reduced fertility but also as a result of a separate maternal effect on seed growth. It was proposed that integument-led seed size enhancement was not a direct result of extra cells in the seed coat but was primarily due to an enlarged seed cavity which allowed greater endosperm growth (Schruff *et al*., 2006). Alternatively, increased nutrient uptake in *arf2* mutant seeds could result from an enlarged area of contact between the seed coat and endosperm.

Consequently *arf2* mutants, if fully fertile, would provide a useful tool to test whether an integument-led seed size increase results in enhanced seed yield. However, like many other single gene mutations that alter seed size, *arf2* causes pleiotropic effects throughout development, such as low fertility. Poor self-fertility in *arf2* mutants severely reduces seed yield, thus it is not meaningful to compare the total yield of mutant and wild-type plants. This chapter describes experiments that attempt to circumvent the problem of reduced fertility in *arf2* mutants to enable meaningful yield assessment of integument-led seed growth. In order to provide a realistic assessment of seed yield, the HI of *arf2* mutants was compared with that of wild-type plants.

4.1.3 Strategies to test integument-led yield improvement

1. Semi-dominant heterozygotes. One reportedly successful strategy to avoid the confounding effect of reduced fertility on yield measurements is to use heterozygous mutant plants (Jofuku *et al*., 2005), provided that the mutation is not fully recessive. As some aspects of the *arf2* phenotype behave in a semi-dominant manner (Okushima *et al*., 2005), the seed size and yield of *arf2-9* heterozygotes was determined and compared to wild-type.

2. Restoration of flower opening in *arf2-9* mutants. This transgenic approach to increase seed size and yield in the *arf2* model aimed to improve fertility by restoration of flower opening. We sought to re-introduce wild-type *ARF2* gene expression to specific floral organs using a construct consisting of the putative *APETALA1* (*AP1*) promoter transcriptionally fused to the *ARF2* coding region (*pAP1::ARF2*). *AP1* is expressed throughout young floral primordia, but later transcripts become localised to the sepals and petals, and are not found elsewhere in the plant (Mandel *et al*., 1992). Therefore, this approach should enable flower opening by restoring sepal elongation to that of wild-type whilst retaining the *arf2* large-seed phenotype. The *pAP1::ARF2* vector was previously cloned and transformed into *arf2-9* plants (Schruff, 2006); this chapter describes the identification of plants homozygous for a single copy of the transgene and a detailed analysis of their phenotype.

3. Increasing seed size through targeted suppression of *ARF2*. A further transgenic approach sought to increase seed size in wild-type plants by suppressing *ARF2* expression specifically within seeds through RNAi-mediated gene silencing. This approach is more appropriate for applying to crop species as it does not require the use of a mutant background as in the case of the *pAP1::ARF2* strategy described above, and consequently the plants should be fully fertile. Two promoters were selected for this strategy. The *INNER NO OUTER* (*INO*) promoter drives gene expression solely in the abaxial outer integument (Villanueva *et al*., 1999) and has been used to generate a *pINO::ARF2 RNAi* construct (Schruff, 2006). In contrast, the *SHATTERPROOF 2* (*SHP2*) promoter directs gene expression throughout the carpel prior to cell type differentiation and later gene expression is confined to the ovules and valve margins (Savidge *et al*., 1995; Dinneny *et al*., 2005). This chapter describes the cloning of the *pSHP2::ARF2 RNAi* construct and the effect of both the *pINO::SHP2 RNAi* and *pSHP2::ARF2 RNAi* transgenes *in planta*.

4.2 Results

4.2.1 Yield assessment of heterozygous arf2-9 plants

Previous analysis of heterozygous plants had shown that the *arf2* mutant allele plays a recessive role in floral morphology, so that the plants are fully fertile; however, it is semi-dominant for some aspects of the phenotype, including the stem diameter and plant height (Okushima *et al*., 2005). It is therefore possible that *arf2* heterozygotes may also retain increased seed size and, in addition, show a higher yield than wild-type. Seed size and yield was compared in restricted and unrestricted pollinations (Section 3.2.3). A recent study suggested that genetic reduction of the *AP2* copy number may enhance yield, as heterozygous *ap2-10* mutants have been reported to produce larger seeds and a greater total seed weight than the wild-type (Jofuku *et al*., 2005). The seed size and yield of *ap2-7* and *ap2-11* heterozygotes was therefore determined in order to compare the results with *arf2-9* heterozygotes. The *ap2-7*, *ap2-10* and *ap2-11* alleles are all considered to have strong homozygous mutant phenotypes (Jofuku *et al*., 2005, Ohto *et al*., 2005).

As reported previously (Jofuku *et al*., 2005; Ohto *et al*., 2005; Schruff *et al*., 2005), homozygous *arf2-9*, *ap2-7* and *ap2-11* mutants produced heavier seeds than the wild-type, and this was true for both restricted and unrestricted pollinations (Table 4.1). Following restricted pollination, the seeds of *arf2-9* heterozygotes were 9.9% heavier than those of Col-3, despite containing the same number of seeds per silique, and this difference was significant. In contrast, the seeds of *ap2-7* and *ap2-11* heterozygotes were 4.7% and 5.0% respectively, smaller than those of Col-0 in restricted pollination. In unrestricted pollination, the mean seed weight of *arf2-9*, *ap2-7* and *ap2-11* heterozygotes was not significantly different from that of the wild-type. In addition, the total seed yield of *arf2-9*, *ap2-7* and *ap2-11* heterozygotes was not significantly different from that of the wild-type. In summary, although *arf2-9* heterozygotes produced larger seeds than the wild-type in restricted pollinations, the total seed yield was not increased by reducing the copy number of *ARF2*. Therefore an alternative strategy was required in order to conduct meaningful trials on the increased seed size and yield in *arf2* mutants.

Table 4.1: Seed weight and yield analysis of *arf2-9, ap2-7 and ap2-11* homozygous and heterozygous mutants compared with wild-type.

	Restricted Pollinations		Unrestricted Pollinations	
Genotype	**Mean Seed Weight (µg)**	**Seed Number per Silique**	**Mean Seed Weight (µg)**	**Total Seed Yield (g)**
Col-3	34.4 ± 0.4^{a}	60 ± 2.0	20.3 ± 0.4^{a}	1.18 ± 0.05^{a}
arf2-9	39.7 ± 0.6^{b}	47 ± 2.1	36.3 ± 0.7^{b}	0.05 ± 0.01^{b}
arf2-9 heterozygote	37.8 ± 0.4^{c}	61 ± 1.4	21.8 ± 0.3^{a}	0.96 ± 0.08^{a}
Col-0	34.3 ± 0.3^{a}	59 ± 1.7	20.6 ± 0.3^{a}	1.12 ± 0.08^{a}
ap2-7	37.6 ± 0.4^{b}	69 ± 1.8	35.8 ± 0.9^{b}	0.02 ± 0.00^{b}
ap2-7 heterozygote	32.7 ± 0.2^{c}	63 ± 1.9	21.0 ± 0.3^{a}	1.12 ± 0.03^{a}
ap2-11	38.4 ± 0.3^{b}	71 ± 1.8	37.1 ± 0.9^{b}	0.01 ± 0.00^{b}
ap2-11 heterozygote	32.6 ± 0.4^{c}	71 ± 1.6	21.0 ± 0.3^{a}	1.10 ± 0.03^{a}

Mean ± s.e.m. $n \geq 7$. For each set of comparisons, seed weight and yield values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. *arf2-9* mutants were compared to Col-3 whereas *ap2-7* and *ap2-11* mutants were compared to Col-0. Homozygous *arf2-9*, *ap2-7* and *ap2-11* mutants were manually pollinated in restricted pollinations in order to achieve full seed set.

4.2.2 pAP1::ARF2 transgene restores flower opening in arf2-9 mutants

In order to restore flower opening and improve fertility, wild-type *ARF2* gene expression was re-introduced into the sepals and petals of *arf2-9* mutants using a construct consisting of the putative *AP1* promoter transcriptionally fused to the *ARF2* coding region (*pAP1::ARF2*). The *pAP1::ARF2* vector was previously designed, cloned and transformed into *Arabidopsis* Col-0 plants (Schruff, 2006). The expression cassette consists of a 1.7 kb sequence upstream of the *AP1* coding region transcriptionally fused to the *ARF2* coding region, along with an ocs 3' terminator. *arf2-9* mutants carrying the *pAP1::ARF2* construct were selected on kanamycin-containing media and confirmed by PCR with primers 'AP1 F check' and 'MNT R check' (Primers and PCR conditions are detailed in Table 2.1 and Table 2.2). From this point onwards, *arf2-9* mutants containing the *pAP1::ARF2* construct are referred to as *pAP1::ARF2* plants.

Introduction of the *pAP1::ARF2* construct into *arf2-9* mutants restored flower opening (Figure 4.1a,b,c). Five independent transformation lines displaying open flowers were analysed for the presence of a single copy of the transgene, high seed weight and good fertility, as judged by the number of elongated siliques (Figure 4.1d). Homozygous

T_3 plants derived from the transformation line which met all the above criteria, *pAP1::ARF2* -5, were used for detailed phenotype analysis.

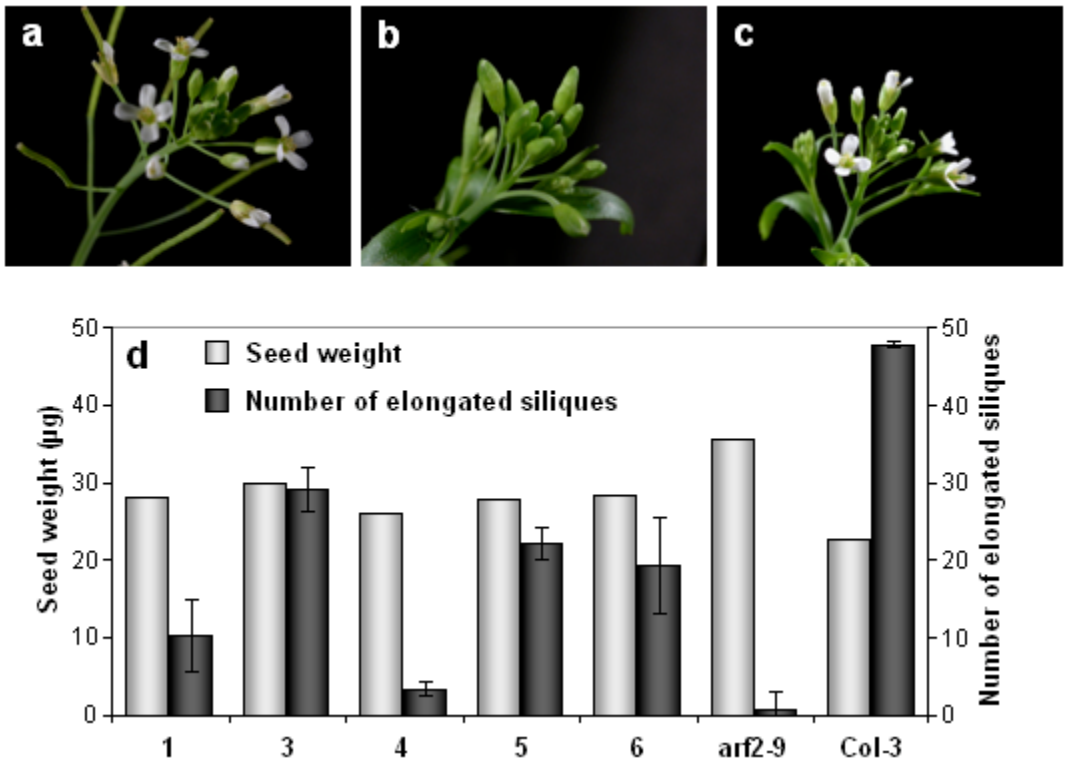

Figure 4.1: The *pAP1::ARF2* construct restores wild-type flower opening to *arf2-9* mutants. Primary inflorescence of (a) Col-3, (b) *arf2-9* and (c) *pAP1::ARF2* plants. (d) Preliminary phenotype analysis of five independent *pAP1::ARF2* transformation lines. Mean seed weight analysis was carried out on T_2 generation seeds. n = 1 batch of 50 seeds. Following the termination of flowering, the number of elongated siliques on the primary inflorescence only was determined on T_2 plants. n = 10 plants. Error bars = s.e.m.

To understand how flower opening was achieved by introduction of the *pAP1::ARF2* construct, the lengths of the floral organs in stage 13 flowers (anthers dehisced, receptive stigma) (Smyth *et al*., 1990) at position 25 on the primary inflorescence were measured. In *pAP1::ARF2* plants, sepals and petals were the same length as the wild-type organs, whereas the gynoecia (in which *AP1* is not expressed) remained over-elongated (Figure 4.2a). Stamen elongation was increased in *pAP1::ARF2* plants compared with *arf2-9* mutants possibly due to an indirect effect of flower opening that removed the physical obstruction of closed sepals.

Although *arf2* mutant flowers failed to open and self-fertilise during the majority of the lifecycle, late flowers did set seed (Schruff *et al*., 2006). Consistent with this

observation, early *arf2-9* flowers had a petal to sepal ratio of less than unity, which was associated with a failure of buds to open, whereas, in late flowers, the petals extended slightly longer than the sepals, allowing partial flower opening. Unlike in *arf2-9* mutants, *pAP1::ARF2* plants had a petal to sepal ratio of greater than unity throughout the lifecycle (Figure 4.2b), indicating that all flowers were open.

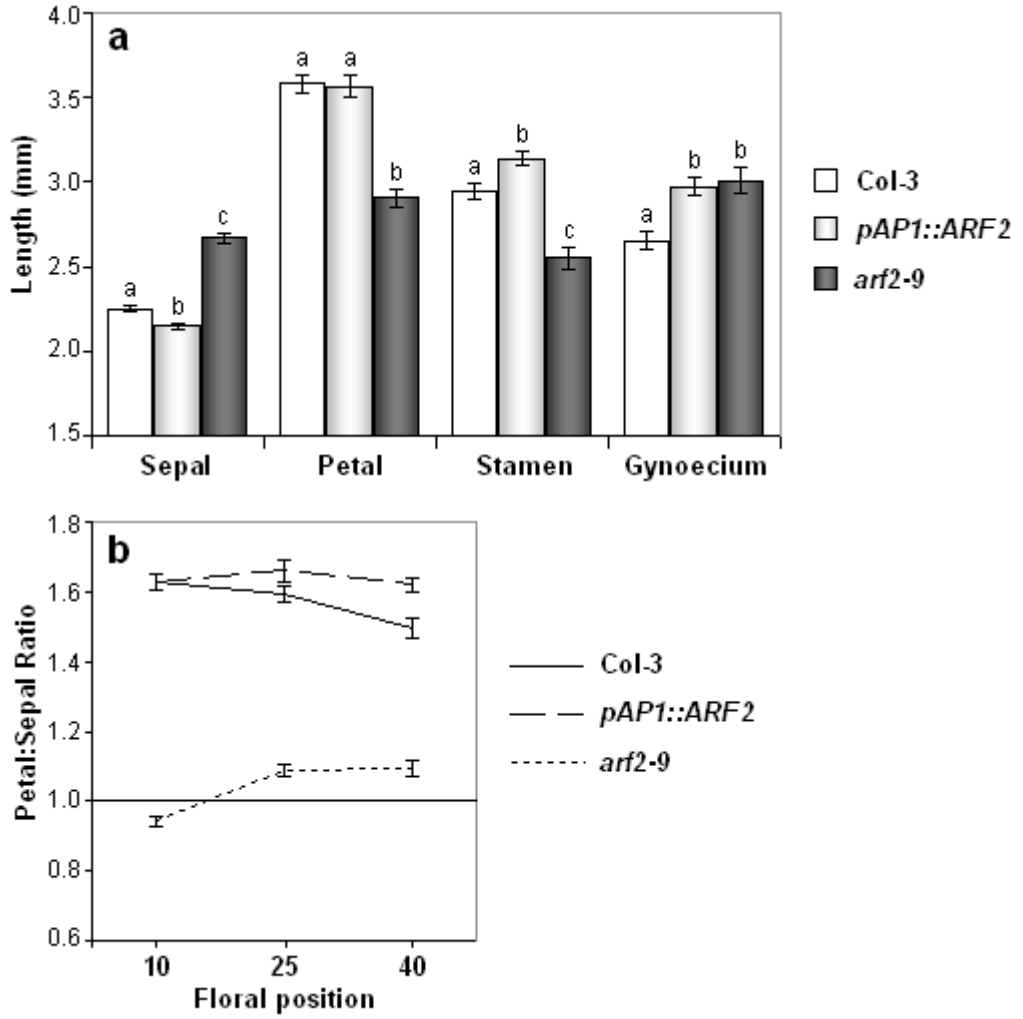

Figure 4.2: The *pAP1::ARF2* construct restores sepal and petal growth allowing flower opening in *arf2-9* mutants.

(a) Lengths of floral organs in stage 13 flowers (Smyth *et al*., 1990) at position 25 on the primary inflorescence. For each set of comparisons, values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. (b) Petal to sepal ratio at three positions on the primary inflorescence: flowers 10, 25 and 40. A petal to sepal ratio of greater than unity represents the ability of flowers to open. Error bars = s.e.m. n = 10 flowers for each genotype.

Although *pAP1::ARF2* plants had open flowers, they retained most features of the *arf2-9* homozygous mutant phenotype, including late flowering, thickened inflorescence stems and increased side shoot number (Table 4.2 and Figure 4.3). In contrast, most features of *arf2-9* heterozygotes were similar to those of Col-3.

Table 4.2: Life history traits of wild-type, *arf2-9* homozygous and heterozygous mutant, and *pAP1::ARF2* plants.

Genotype:	**Col-3**	***pAP1::ARF2***	***arf2-9***	***arf2-9* heterozygote**
Flowering time (days)	34.5 ± 0.17	43.4 ± 0.31	43.8 ± 0.13	35.7 ± 0.54
Stem thickness (cm)	1.4 ± 0.04	1.9 ± 0.04	1.8 ± 0.04	1.5 ± 0.05
Side shoot number	7.9 ± 0.1	10.3 ± 0.2	11.0 ± 0.3	8.0 ± 0.3
Plant height (cm)	44.9 ± 1.82	51.8 ± 0.41	58.3 ± 1.10	52.1 ± 0.57
Silique length (mm)	14.7 ± 0.25	17.7 ± 0.80	7.7 ± 1.33	15.8 ± 0.14
Ovule number per silique	68.0 ± 1.15	64.6 ± 0.95	60.9 ± 1.37	68.3 ± 1.92
Flower number	58.7 ± 0.98	62.1 ± 1.12	59.5 ± 1.15	56.9 ± 0.69
Elongated silique number	58.7 ± 0.98	54.4 ± 0.93	17.6 ± 1.80	53.8 ± 0.90

Mean ± s.e.m. n = 10 plants per genotype. Flowering time scored as the period from germination to opening of the first flower. Stem thickness, i.e. the diameter of the primary inflorescence, was measured between the first and second nodes from the base when the stems were ~30 cm in length. Ovule number per silique was determined in the 20th silique on the primary inflorescence at 3 DAP. Silique length, plant height, flower number and side shoot number were all determined after flowering had terminated. Silique length was determined in the 30th silique on the primary inflorescence. Side shoot number refers to the number of branches on the primary inflorescence plus the number of additional bolts from the rosette. The numbers of flowers and elongated siliques refers to only the primary inflorescence.

Figure 4.3: *pAP1::ARF2* plants retain many features of *arf2-9* mutants.

Five week old plants of Col-3, *arf2-9* and *pAP1::ARF2* demonstrating the similar inflorescence architecture, flowering time and stem thickness of *pAP1::ARF2* and *arf2-9* plants.

4.2.3 pAP1::ARF2 plants retain the arf2-9 seed phenotype

While most features of the *arf2-9* phenotype were retained, it was important to assess seed development in *pAP1::ARF2* plants to determine whether integument cell over-proliferation, increased seed size and the distinctive seed shape were also preserved. Homozygous *arf2-9* mutants have large pointed seeds as a result of extra cell divisions in the ovule integuments prior to fertilisation (Schruff *et al*., 2006). Developing and mature seeds of *pAP1::ARF2* plants were also larger and more pointed than those of the wild-type control, whereas seeds produced by *arf2-9* heterozygotes resembled those of the wild-type (Figure 4.4a-e). Embryogenesis in *arf2-9* mutants is delayed, for example at 5 DAP wild-type seeds reach the heart stage whereas *arf2-9* mutant seeds contain globular to early heart stage embryos (Schruff *et al*., 2006). Consistent with this, *pAP1::ARF2* developing seeds also showed slow embryo growth. In contrast, embryogenesis occurred at the same rate in *arf2-9* heterozygotes as in wild-type seeds.

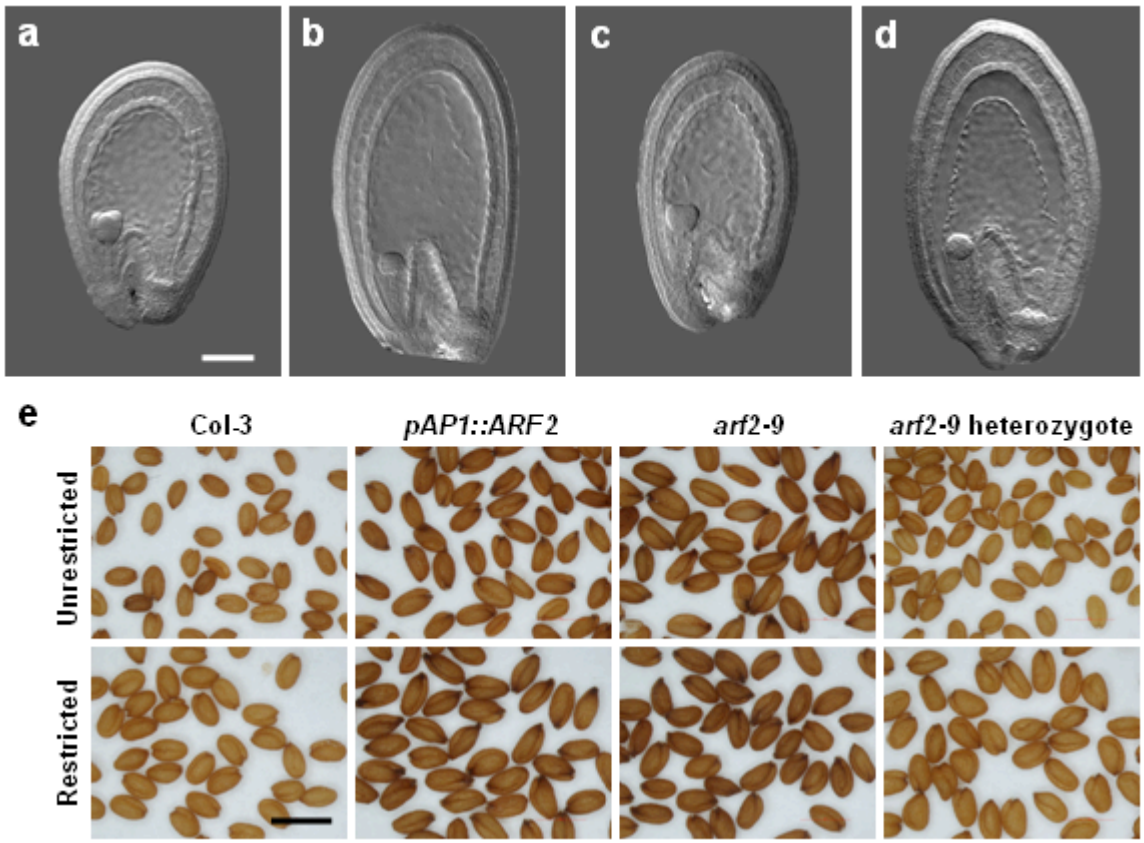

Figure 4.4: Seed development in wild-type, mutant and transgenic plants.

(a-d) Cleared seeds imaged with differential contrast optics at 4 DAP from (a) Col-3, (b) *arf2-9* homozygous, (c) *arf2-9* heterozygous and (d) *pAP1::ARF2* plants during unrestricted pollinations (bar = 100 µm). (e) Mature seeds from restricted and unrestricted pollinations (bar = 1 mm).

To determine whether differences in seed size and shape were related to the numbers of cells in the seed coat post-fertilisation, four seed coat layers were examined: ii1, ii1', oi1, and oi2 (Schneitz *et al*., 1995; Beeckman *et al*., 2000). At 2-5 DAP, all of these layers in the developing seeds of *arf2-9* homozygous mutants, with or without the *pAP1::ARF2* construct, contained significantly more cells than the wild-type (Figure 4.5). In addition there was no significant difference between the numbers of cells in each seed coat layer of the *arf2-9* homozygous mutants and *pAP1::ARF2* transgenic plants. In *arf2-9* heterozygotes, although some seed coat layers contained more cells than the wild-type at some time points, there was no consistent increase. In summary, the *arf2* mutant seed phenotype is preserved in *arf2-9* plants carrying the *pAP1::ARF2* transgene.

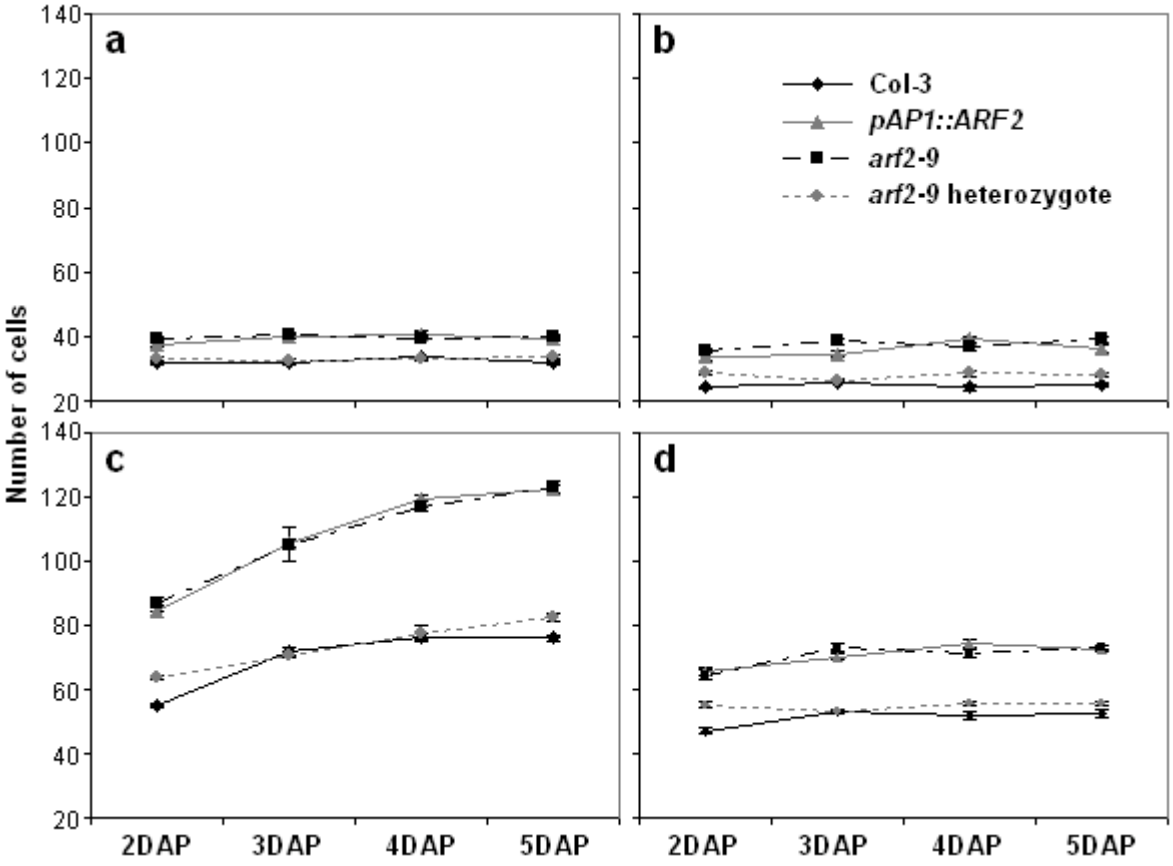

Figure 4.5: Integument cell number in wild-type, mutant and transgenic seeds.

Number of cells in (a) ii1, (b) ii1', (c) oi1, (d) oi2 layers of the seed coat of wild-type, *arf2-9* homozygous and heterozygous, and *pAP1::ARF2* developing seeds at 2-5 DAP. Error bars = s.e.m. n = 10.

4.2.4 Assessing the fertility, seed size and yield of pAP1::ARF2 plants

pAP1::ARF2 plants were assessed using both restricted and unrestricted pollinations (Figure 4.6a) to establish whether over-sized integuments conferred an increased seed weight. In both treatments, *pAP1::ARF2* seeds were larger than those of the wild-type. *pAP1::ARF2* seeds from unrestricted pollination were 35.2% heavier than those of the wild-type, whereas *arf2-9* homozygous mutant seeds were 78.8% heavier. Previously it was proposed that increased seed size in *arf2-9* homozygotes has two components: direct effects of the mutation on seed development, and an indirect effect caused by poor self-fertility (Schruff *et al*., 2006). Therefore, the reduced seed weight of *pAP1::ARF2* relative to *arf2-9* was expected, as the transgenic plants were more fertile. In restricted pollination, although holding the seed set constant narrowed the gap between mutant and wild-type seed weights, *pAP1::ARF2* seeds remained heavier than those of the wild-type, and this difference was significant (Student's *t*-test; Col-3 vs. *pAP1::ARF2*, $P < 0.001$). The mean number of seeds per silique in restricted pollination was 59.7 ± 2.04 (n = 10 siliques) for

the wild-type and 58.1 ± 0.46 (n = 8) for *pAP1::ARF2*; therefore the increase in seed weight did not result from a decrease in seed number.

While flower opening had been restored in *arf2-9* plants containing the *pAP1::ARF2* transgene, it was necessary to establish whether self-pollination was achieved. Poor fertility in *arf2-9* mutants is not only associated with the failure of floral bud opening, but also with the over-elongated gynoecia. Unlike *arf2-9* mutants, which have a stamen to gynoecium ratio of less than unity for most of their lifecycle, *pAP1::ARF2* plants have a slightly higher ratio (Figure 4.6b), as the stamens and gynoecia are both over-elongated (Figure 4.2a). Thus flower opening in *pAP1::ARF2* plants allowed stamen elongation, which brings pollen into the vicinity of the stigma on the enlarged gynoecium.

Since *pAP1::ARF2* plants produce flowers that self-pollinate and enlarged seeds, it was possible to conduct a meaningful yield assessment. The total seed yields of wild-type plants and *arf2-9* mutants with and without the *pAP1::ARF2* construct were measured (Figure 4.6c). Compared with the mutant alone, the total seed yield of *pAP1::ARF2* plants increased more than 15-fold. This corresponded to an increase in seed number per plant from a mean of 1,316 ± 186 (n = 7) in *arf2-9* to 27,110 ± 3,042 (n = 8) on *pAP1::ARF2*. However, the yield of *pAP1::ARF2* plants was lower than that of the wild-type, consistent with the production of fewer seeds per plant (57,786 ± 2,182 seeds per plant in Col-3, n = 8). In order to investigate this loss of yield, the percentage of fertilised ovules in the 20th silique on the primary inflorescence was calculated (Figure 4.6d). Seed set was increased in *arf2-9* mutants from 20.5% to 78.2% by introduction of the *pAP1::ARF2* construct, but fewer ovules were fertilised than in the wild-type. Although *pAP1::ARF2* flowers usually have a stamen and gynoecium ratio of greater than unity, it is not as high as in the wild-type (Figure 4.6b), and it is probable that an insufficient number of pollen grains are shed on to the stigma to give full fertility.

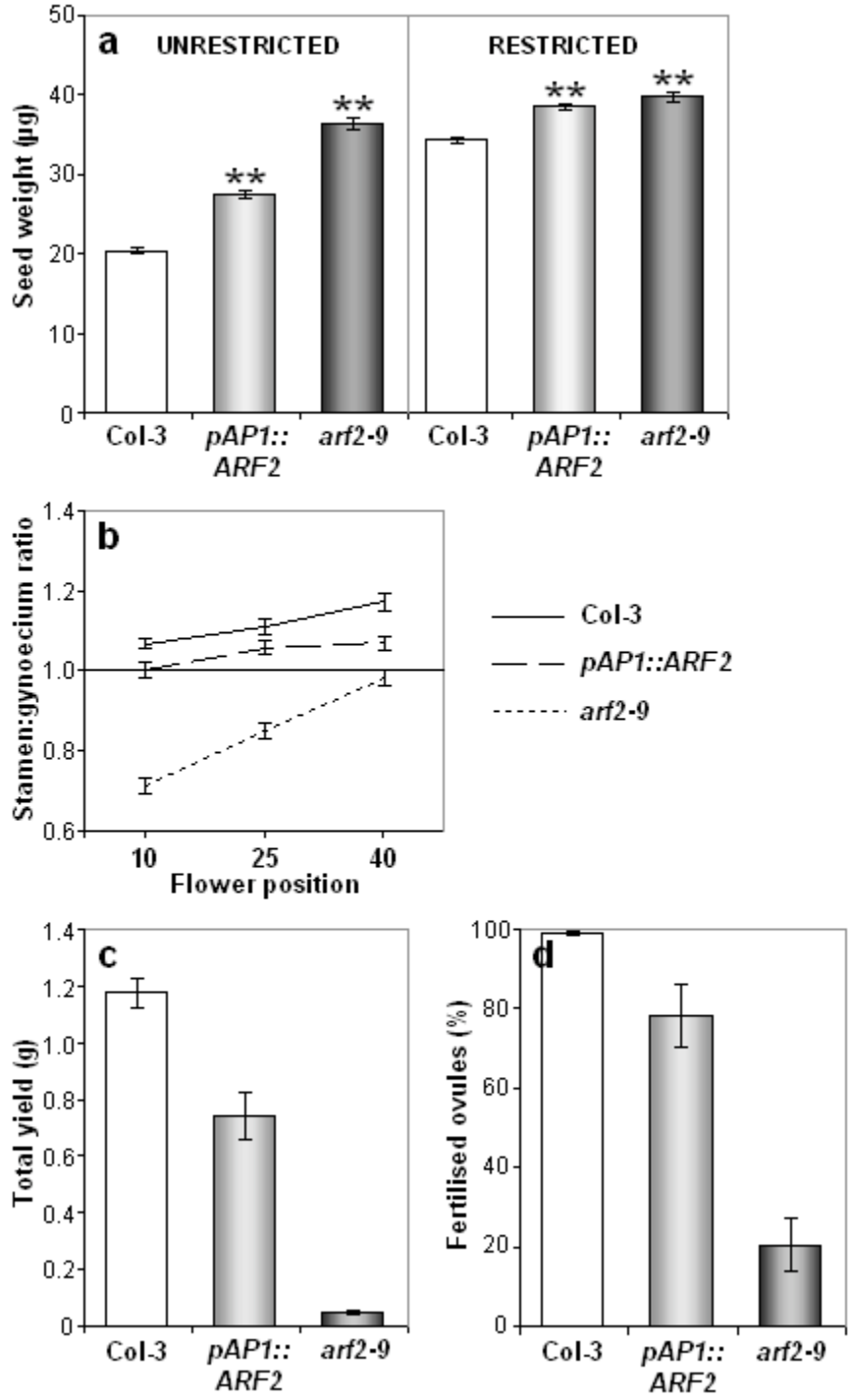

Figure 4.6: *pAP1::ARF2* plants produce larger seeds than the wild-type and show greater fertility and yield than *arf2-9* mutants.

(a) Comparison of seed weights following unrestricted (left) and restricted (right) pollinations. For each set of comparisons, values that differ at the 0.01 significance level (Student's *t*-test) from wild-type are labelled with **. (b) Stamen to gynoecium ratio at three positions on the primary inflorescence: flowers 10, 25 and 40. A stamen to gynoecium ratio of greater than unity is required to deposit pollen onto the stigma. (c) Yield analysis: total seed weight per plant. (d) Percentage of ovules that were fertilised in the 20th silique on the primary inflorescence. Error bars = s.e.m. $n \geq 7$.

4.2.5 Fatty acid content of pAP1::ARF2 and arf2-9 seeds

Oil is a key seed storage compound with high economic value for food, feed and industrial markets. *Arabidopsis* is a close relative of a major oilseed crop, *Brassica napus* (oilseed rape), and therefore provides a model to study oilseeds. The total fatty acid content of wild-type, *pAP1::ARF2* and *arf2-9* seeds were compared in order to determine if enlarged seed size was associated with increased oil content (Table 4.3) (lipid analysis was performed by Tony Larson and Ian Graham, University of York). Following restricted seed set, *arf2-9* mutant and *pAP1::ARF2* seeds contained 28% and 42% more fatty acid respectively than wild-type seeds. However, there was no significant differences in the total fatty acid content per microgram of seed (ANOVA followed by Tukey's multiple comparisons, P = 0.592). Therefore, *arf2-9* and *pAP1::ARF2* seeds showed increased lipid content simply because they were larger. Similar results are shown for seeds produced in unrestricted pollinations, however in contrast the total fatty acid per microgram of seed was lower in *arf2-9* and *pAP1::ARF2* seeds than found in the wild-type.

Table 4.3: Lipid content of wild-type, *arf2-9* homozygous and *pAP1::ARF2* seeds.

	Genotype	Total fatty acid (µg/seed)	Total fatty acid (µg/µg seed)
Restricted	Col-3	8.84 ± 0.63^{a}	0.27 ± 0.02^{a}
	pAP1::ARF2	12.56 ± 0.42 (+42%)^{b}	0.30 ± 0.02 (+10%)^{a}
	arf2-9	11.31 ± 0.89 (+28%)^{a}	0.28 ± 0.02 (+3%)^{a}
Unrestricted	Col-3	8.44 ± 0.24^{a}	0.40 ± 0.01^{a}
	pAP1::ARF2	9.00 ± 0.32 (+7%)^{a}	0.34 ± 0.01 (-15%)^{a}
	arf2-9	11.33 ± 0.76 (+34%)^{b}	0.33 ± 0.03 (-17.5%)^{b}

Mean ± s.e.m. n = 5 batches of 15 seeds. Percentage difference from wild-type is given in parentheses. Separate statistical analyses were carried out for seeds produced in restricted and unrestricted pollinations. For each set of comparisons, values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different superscript letters. Lipid analysis was performed by Tony Larson and Ian Graham, University of York.

The maintenance of consistent fatty acid ratios may be important if this technology is transferred to oilseed crops, therefore the fatty acid composition of wild-type, mutant and transgenic seeds from restricted and unrestricted pollinations was also analysed. Following restricted pollination, the *arf2-9* mutation, with or without *pAP1::ARF2*, did not cause dramatic changes to the fatty acid profile compared with wild-type seeds. For instance, no significant difference was recorded in the level of linoleic acid (18:2n6c), the most abundant fatty acid in *Arabidopsis*. However, in *pAP1::ARF2* and *arf2-9* seeds, several fatty acids showed levels that were significantly different from those of wild-type (Figure 4.7a). Similar results were obtained for seeds produced in unrestricted pollinations (Figure 4.7b).

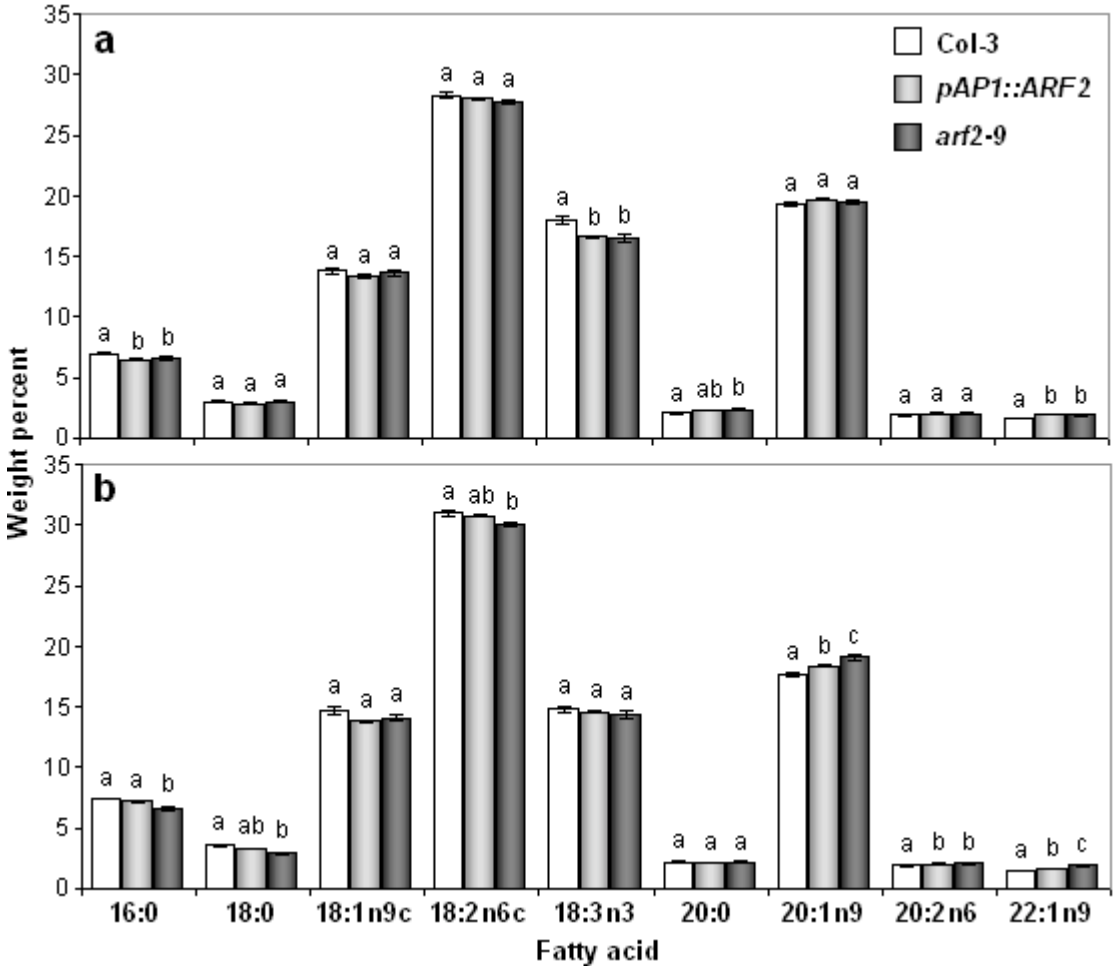

Figure 4.7: Analysis of fatty acid composition in wild-type, mutant and transgenic seeds.

Fatty acid composition of wild-type, *arf2-9* homozygous and *pAP1::ARF2* seeds in (a) restricted and (b) unrestricted pollinations. For each set of comparisons, values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different superscript letters. Error bars = s.e.m. n = 5 batches of 15 seeds. Lipid analysis was performed by Tony Larson and Ian Graham, University of York.

4.2.6 pAP1::ARF2 transgene increases the HI of arf2-9 mutants

In order to provide a realistic assessment of seed yield, the HI of *pAP1::ARF2* and other plants was calculated as: seed yield / (dry mass of aerial organs + seed yield). For each genotype tested, the seed yield is shown in Figure 4.8a (also in Table 4.1 and Figure 4.6c), dry mass of aerial organs in Figure 4.8b, total dry mass (i.e. mass of aerial organs plus seed yield) in Figure 4.8c, and HI in Figure 4.8d.

Homozygous *arf2-9* mutants have a large dry mass (Figure 4.8b), consistent with previous reports that *arf2* mutations cause the over-growth of many organs (Ellis *et al*., 2005, Okushima *et al*., 2005, Schruff *et al*., 2006). The combined effect of the large mass and poor self-fertility of the *arf2-9* mutants results in a low HI (Figure 4.8d). *pAP1::ARF2* plants have a smaller dry mass of aerial organs and a greater seed yield than *arf2-9* mutants. Hence, as expected, the introduction of the *pAP1::ARF2* construct substantially improves the HI from 0.01 to 0.195. Col-3 plants have an even higher HI, as they produce less mass and a greater yield than *arf2-9* and *pAP1::ARF2* plants. Similar to *arf2-9* mutants, *ap2-7* and *ap2-11* mutants have a low HI because of poor self-fertility. Although there has been no report of increased cell division or organ size *in ap2* mutants, apart from that in seeds, this analysis revealed an increase in dry mass of aerial organs relative to that of the wild-type (Student's *t*-test; *ap2-7* vs. Col-0, $P \leq 0.001$; *ap2-11* vs. Col-0, $P = 0.009$).

The seed yields of heterozygous *arf2-9*, *ap2-7* and *ap2-11* plants showed no significant difference from that of the wild-type (Table 4.1, Figure 4.8a), and there was likewise no difference in the dry mass of aerial organs (Student's *t*-test; *arf2*-9 heterozygote vs. Col-3, $P = 0.125$; *ap2-7* heterozygote vs. Col-0, $P = 0.175$; *ap2-11* heterozygote vs. Col-0, $P = 0.453$). Consistent with this, the HI of *ap2-11* heterozygotes was not significantly different from that of the wild-type (Student's *t*-test; *ap2-11* heterozygote vs. Col-0, $P = 0.446$). However, the HI of *ap2-7* heterozygotes was significantly higher than the wild-type (Student's *t*-test; *ap2-7* heterozygote vs. Col-0, $P = 0.005$) and the HI of *arf2-9* heterozygotes was significantly lower than that of the wild-type (Student's *t*-test; *arf2-9* heterozygote vs. Col-3, $P = 0.0023$).

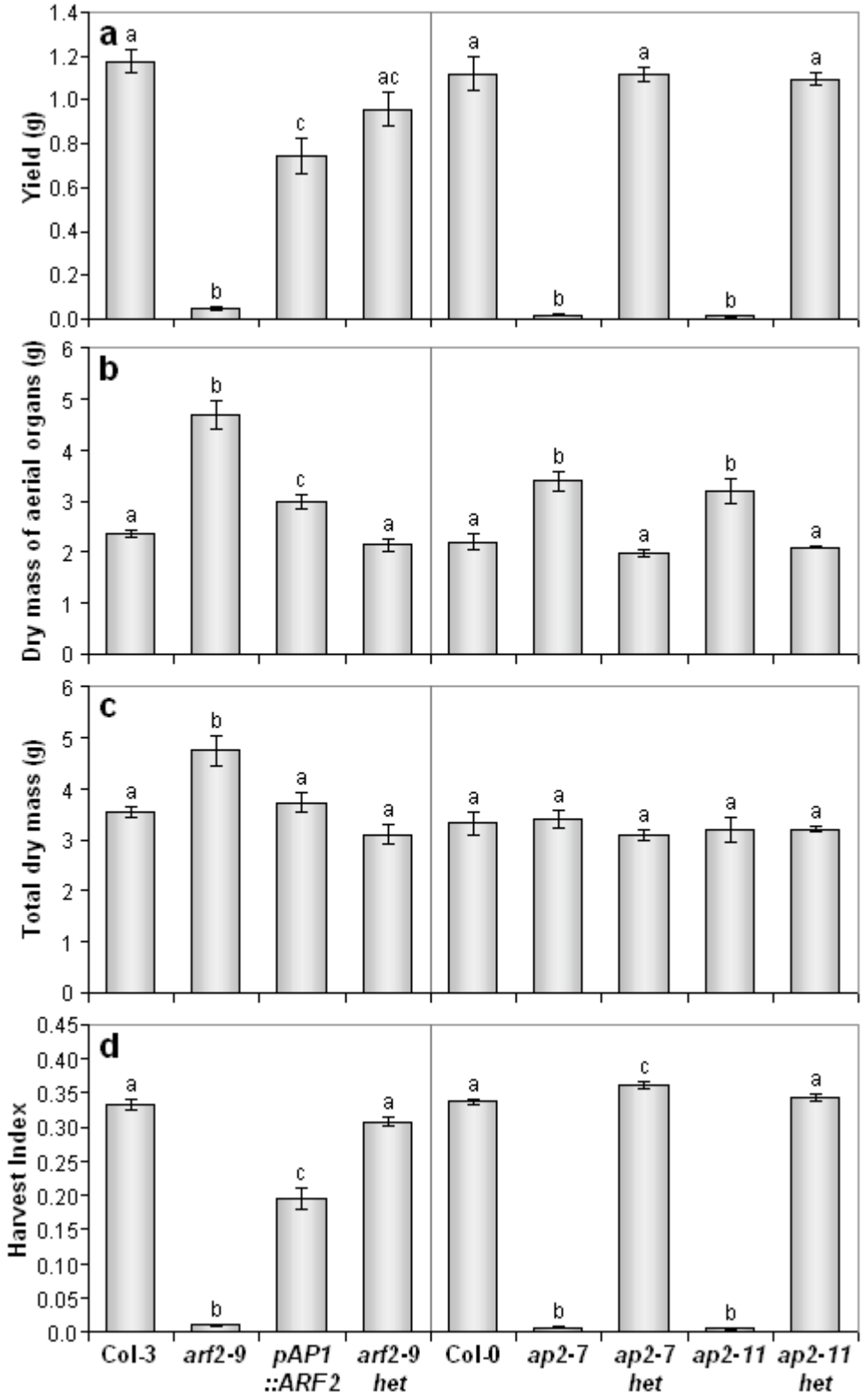

Figure 4.8: HI measurement in wild-type, mutant and transgenic plants.

(a) Seed yield. (b) Dry mass of aerial organs. (c) Total dry mass (seed yield plus dry mass of aerial organs). (d) HI. For each set of comparisons, seed weight and yield values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. arf2-9 mutants and *pAP1::ARF2* plants were compared to Col-3 (left) whereas ap2-7 and ap2-11 mutants were compared to Col-0 (right). Error bars = s.e.m. $n \geq 6$.

4.2.7 Selection of plants carrying the pINO::ARF2 RNAi transgene

To increase seed size in wild-type plants and maintain full fertility, *ARF2* was suppressed specifically in the ovule integuments. The *INO* promoter confers gene expression solely in the abaxial outer integument (Villanueva *et al*., 1999). The *pINO::ARF2 RNAi* construct was designed, cloned and transformed into *Arabidopsis* Col-0 plants (Schruff, 2006). The expression cassette contains a sequence of 1,770 bp upstream of the *INO* coding region, and two 570 bp *ARF2* fragments, one in the forward and one in the reverse orientation and separated by a CHSA intron. Positive *Arabidopsis* transformants were selected by treatment with BASTA (glufosinate). PCR using primers 'INOMNTRNAi F' and 'INOMNTRNAi R' confirmed that the transgene was present (primers and PCR conditions are detailed in Table 2.1 and Table 2.2). Following this, 13 independent transformation lines were analysed for presence of a single copy of the transgene and mean seed weight (Figure 4.9). Two lines were chosen for further analysis. Both *pINO::ARF2 RNAi*-1 and *pINO::ARF2 RNAi*-3 showed high seed weight, although *pINO::ARF2 RNAi*-1 contained multiple copies of the transgene.

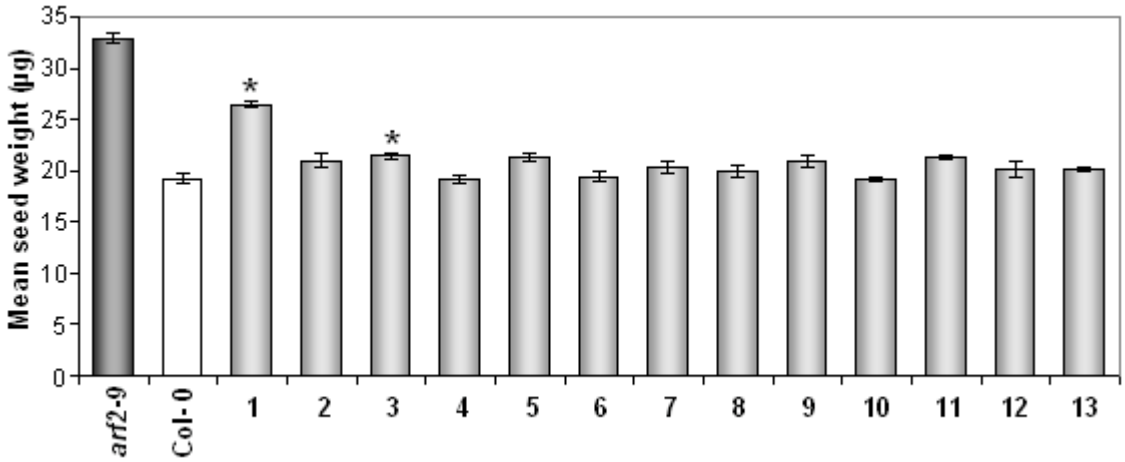

Figure 4.9: Seed weight analysis of T_1 generation *pINO::ARF2 RNAi* plants.

Comparison of the mean seed weight of 13 independent transformation lines containing the *pINO::ARF2 RNAi* transgene with that of wild-type and the *arf2-9* mutant. * indicates lines selected for further analysis. Error bars = s.e.m. n = 3 batches of 50 seeds per plant.

4.2.8 Seed size, yield and HI assessment of pINO::ARF2 RNAi plants

Seed weight, yield and HI of T_3 generation plants containing the *pINO::ARF2 RNAi* transgene were compared to wild-type to assess whether introduction of the construct led to a realistic improvement in seed yield (Table 4.4). Following restricted pollination, both *pINO::ARF2 RNAi* -1 and *pINO::ARF2 RNAi* -3 produced significantly larger seeds than the wild-type (Student's *t*-test; Col-0 vs. line-1, P = 0.001; Col-0 vs. line-3, P < 0.001). Unlike in *arf2-9* mutants, fertility in *pINO::ARF2 RNAi* plants is uncompromised as the number of seeds set per silique is similar to wild-type, without the requirement for manual pollination. Therefore, the increase in seed weight observed in restricted pollinations was not a consequence of decreased seed number. In contrast, no difference was found between the mean seed weight in unrestricted pollinations (ANOVA followed by Tukey's multiple comparisons, P = 0.220) and the total seed yield (ANOVA followed by Tukey's multiple comparisons, P = 0.611) of the *pINO::ARF2 RNAi* lines and wild-type plants. Surprisingly, the HI of both *pINO::ARF2 RNAi* lines was found to be significantly different from wild-type (Student's *t*-test; Col-0 vs. line-1, P < 0.001; Col-0 vs. line-3, P < 0.001). The decrease in HI in the transgenic plants compared to wild type was associated with an increase in the dry mass of aerial organs rather than a decrease in seed yield.

Table 4.4: Seed weight and yield analysis of *pINO::ARF2 RNAi* lines compared to wild-type.

Genotype	Col-0	*pINO::ARF2 RNAi* -1	*pINO::ARF2 RNAi* -3
Restricted seed weight (g)	34.3 ± 0.28^{a}	35.9 ± 0.28^{b}	36.1 ± 0.23^{b}
Seed number per silique	58.8 ± 1.34	61.6 ± 0.91	57.6 ± 2.50
Unrestricted seed weight (g)	20.6 ± 0.25^{a}	20.5 ± 0.27^{a}	21.3 ± 0.42^{a}
Seed yield (g)	1.12 ± 0.08^{a}	1.04 ± 0.06^{a}	1.01 ± 0.10^{a}
Dry mass of aerial organs (g)	2.20 ± 0.15^{a}	2.67 ± 0.06^{ab}	2.80 ± 0.18^{b}
Total biomass (g)	3.32 ± 0.23^{a}	3.70 ± 0.09^{a}	3.81 ± 0.27^{a}
Harvest index	0.34 ± 0.005^{a}	0.28 ± 0.010^{b}	0.26 ± 0.010^{b}

Error bars = s.e.m. $n \geq 6$. For each set of comparisons values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different superscript letters. Seed number per silique was determined in restricted pollinations.

4.2.9 Assembly of the pSHP2::ARF2 RNAi construct and selection of transgenic plants

As an alternative to the *INO* promoter, the *SHP2* promoter was used to suppress *ARF2* in wild-type plants in order to increase integument proliferation, seed size and yield. The *SHP2* promoter drives expression throughout the carpel prior to cell type differentiation, with gene expression later confined to the ovules and valve margins (Savidge *et al*., 1995; Dinneny *et al*., 2005); therefore, the *SHP2* promoter confers a broader expression pattern than the *INO* promoter. The 2.1 kb *SHP2* promoter fragment upstream of the *SHP2* coding region was amplified from genomic DNA using PCR with primers 'At2g42830 SHP2 prom F' and 'At2g42830 SHP2 prom R' (primers and PCR conditions are detailed in Table 2.3 and Table 2.4). Restriction site linkers, *Eco*RI and *Asc*I, were incorporated into the primers to enable directional cloning into the pFGC5941 RNAi vector. Prior to inserting into pFGC5941, the *SHP2* promoter was first inserted into the pGEM-T using TA cloning and checked for sequence integrity (Figure 4.10a). The CaMV 35S promoter was removed from pFGC5941 using *Eco*RI and *Asc*I, and replaced with the *SHP2* promoter (Figure 4.10b,c) which had been lifted from pGEM-T. The 570 bp *ARF2* fragments were obtained from the *pINO::ARF2 RNAi* construct previously cloned (Schruff, 2006). First, the forward *ARF2* fragment bordered by *Asc*I and *Swa*I was digested out of *pINO::ARF2 RNAi* and inserted alongside the *SHP2* promoter in pFGC5941 (Figure 4.10d). Following this, the reverse *ARF2* fragment bordered by *Bam*HI and *Xba*I was removed from *pINO::ARF2 RNAi* and ligated between the CHSA intron and the ocs 3' (Figure 4.10e), completing the cloning of the *pSHP2::ARF2 RNAi* construct.

Arabidopsis Col-0 plants were transformed with the *pSHP2::ARF2 RNAi* construct and positive transformants were selected by treatment with BASTA (glufosinate) as the *pSHP2::ARF2 RNAi* transgene also carries the BAR gene which confers *in planta* resistance to this herbicide. Additionally, two separate PCR reactions were used to check for presence and integrity of the *pSHP2::ARF2 RNAi* transgene in plants using primers 'SHP2 F check' and 'ARF2 R check' and secondly 'INOMNTRNAi F' and 'AVA ocs3' 2R' (primer sequences and PCR conditions are detailed in Table 2.1 and Table 2.2).

The mean seed weight of 23 *pSHP2::ARF2 RNAi* independent transformation lines was determined (T_1 generation plants) (Figure 4.11). Seven lines displaying the highest seed weights were later assayed for the presence of only a single copy of the transgene. *pSHP2::ARF2 RNAi* -5, *pSHP2::ARF2 RNAi* -22 and *pSHP2::ARF2 RNAi* -26 were selected for more detailed seed weight analysis.

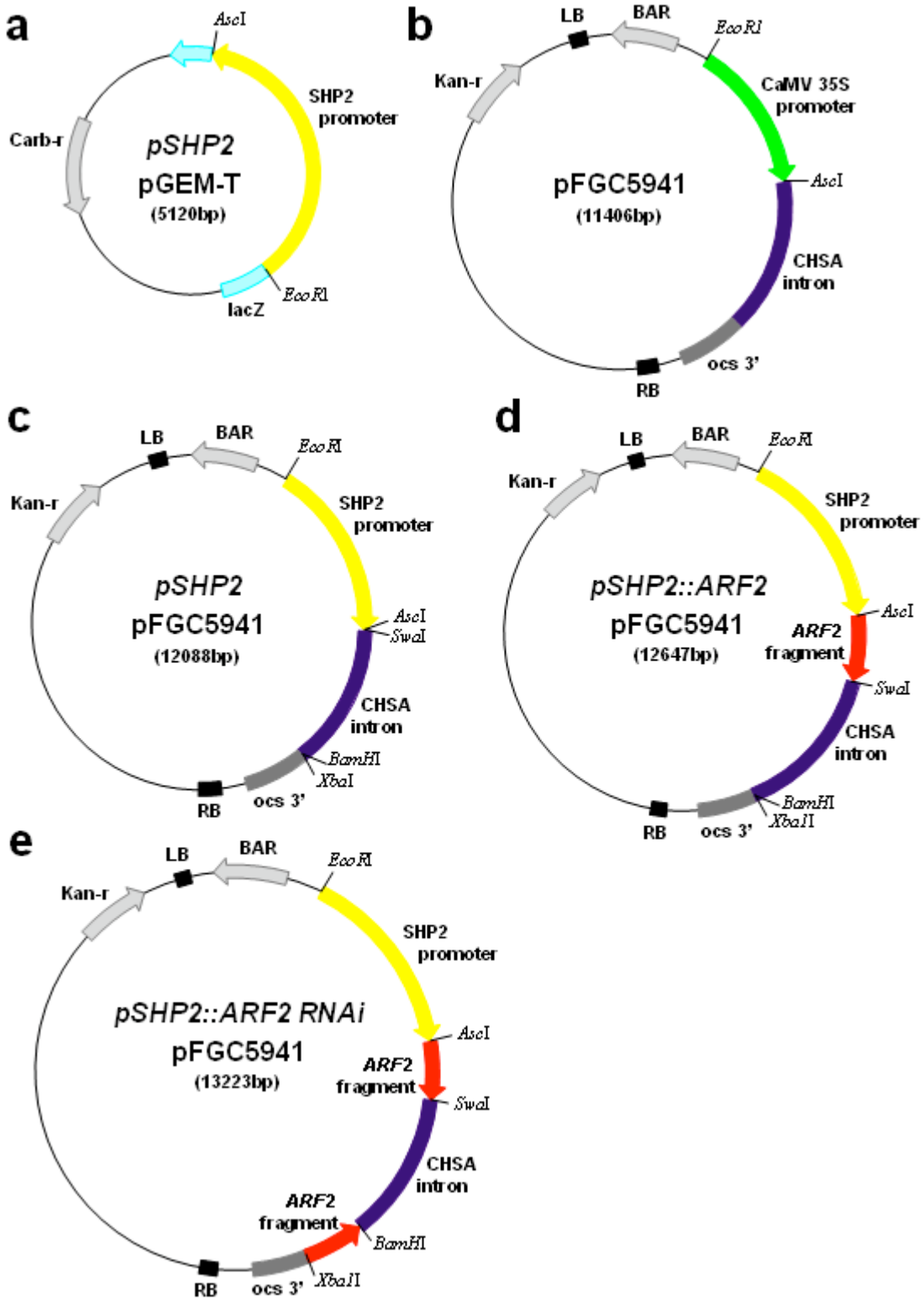

Figure 4.10: Construction of the *pSHP2::ARF2 RNAi* expression vector.

(a) The *SHP2* promoter was first cloned into the pGEM-T vector. (b) The CaMV 35S promoter was removed from pFGC5941 vector using *Eco*RI and *Asc*I. (c) The *SHP2* promoter was ligated between the *Eco*RI and *Asc*I sites of the pFGC5941 vector. (d) The forward *ARF2* fragment was inserted adjacent to the *SHP2* promoter using *Asc*I and *Swa*I. (e) The reverse *ARF2* fragment was inserted between the CHSA intron and the ocs 3' using *Bam*HI and *Xba*I, generating the *pSHP2::ARF2 RNAi* construct.

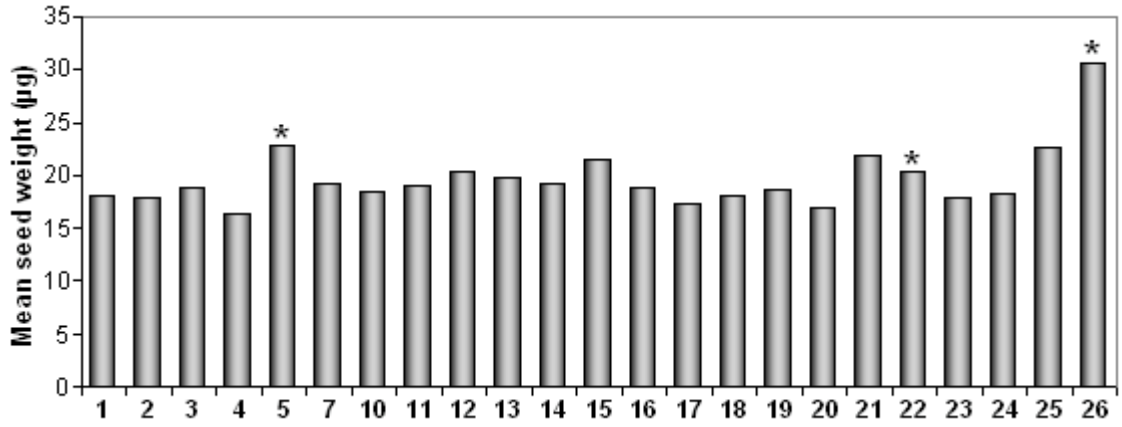

Figure 4.11: Seed weight analysis of T_1 generation *pSHP2::ARF2 RNAi* plants.

Mean seed weight of 23 independent transformation lines of containing the *pSHP2::ARF2 RNAi* transgene. * indicates lines selected for further analysis. Error bars = s.e.m. n = 1 batch of 50 seeds/plant.

4.2.10 Seed size analyses of pSHP2::ARF2 RNAi plants

To determine whether introduction of the *pSHP2::ARF2 RNAi* transgene conferred an increase in seed size, the seed weight of T_3 generation plants from three independent transformation lines was investigated (Figure 4.12a). *pSHP2::ARF2 RNAi* -26 seeds were found to be significantly larger than those of wild-type in both restricted (Student's *t*-test; Col-0 vs. 26, $P < 0.001$) and unrestricted pollinations (Student's *t*-test; Col-0 vs. 26, $P = 0.002$). *pSHP2::ARF2 RNAi* -5 and -22 plants produced significantly larger seeds than wild-type following restricted pollinations (Student's *t*-test; Col-0 vs. 5, $P = 0.001$; Col-0 vs. 22, $P = 0.018$); but were not significantly different from wild-type in unrestricted pollinations.

As plants containing the *pSHP2::ARF2 RNAi* transgene should be fully fertile, restricted pollinations were carried out without manual pollination. To check that seed set was uncompromised the mean number of seeds per silique of all three *pSHP2::ARF2 RNAi* lines in restricted pollinations was determined (Figure 4.12b). This revealed that seed set was reduced by 20.0%, 8.3% and 11.7% in *pSHP2::ARF2 RNAi* -5, *pSHP2::ARF2-RNAi* -22 and *pSHP2::ARF2 RNAi* -2 respectively, compared to the wild-type.

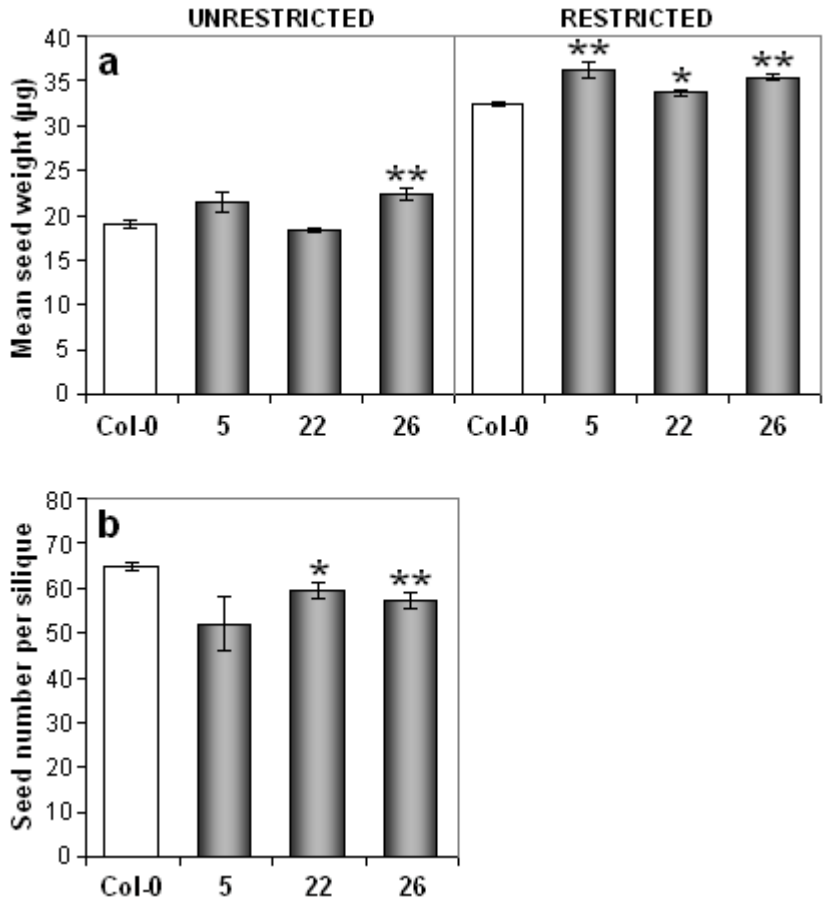

Figure 4.12: *pSHP::ARF2 RNAi* plants produce larger seeds than the wild-type.

(a) Comparison of seed weights of three *pSHP2::ARF2 RNAi* lines with wild-type following unrestricted (left) and restricted (right) pollinations. (b) Analysis of seed set in *pSHP2::ARF2 RNAi* lines, as determined by the seed number per silique in restricted pollinations. For each set of comparisons, values that differ from wild-type at the 0.05 significance level (Student's *t*-test) are labelled with *, and at the 0.01 significance level with **. Error bars = s.e.m. $n \geq 5$ plants.

4.2.11 pSHP2::ARF2 RNAi plants show decreased ARF2 transcript levels

To confirm that the *pSHP2::ARF2 RNAi* transgene was suppressing the *ARF2* gene in developing seeds and siliques, semi-quantitative RT-PCR was carried out to compare the level of *ARF2* transcripts in both transgenic and wild-type plants. The cDNA template was obtained from the reverse transcription of RNA from 3 DAP siliques. Primers within the *ARF2* and *GAPC* genes were designed to bridge an exon-exon junction to ensure that only cDNA, and not any genomic DNA contaminants were amplified (Figure 4.13a). The *ARF2* fragment was amplified using 'ARF2 RT-PCR For' and 'ARF2 RT-PCR Rev', and *GAPC* using 'GAPC Forward' and 'GAPC Reverse' (primer sequences and PCR conditions are detailed in Table 2.1 and Table 2.2). The PCR reveals that *ARF2* transcripts were reduced in both *pSHP2::ARF2 RNAi* -5 and *pSHP2::ARF RNAi* -26 compared to wild-type, as the bands are of lower intensity (Figure 4.13b). However, there was no clear difference between the *ARF2* transcript level of *pSHP2::ARF2 RNAi* -22 and wild-type.

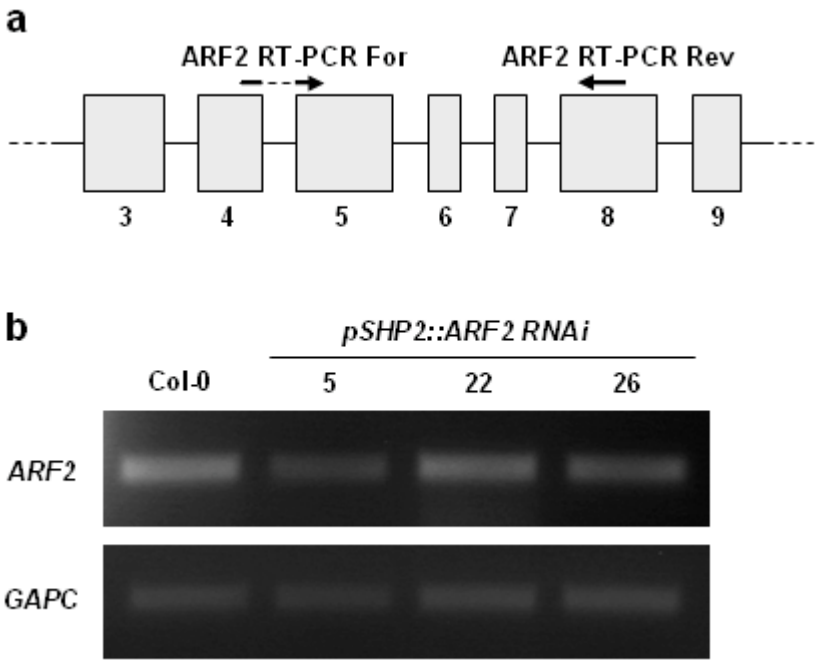

Figure 4.13: *ARF2* transcript levels in *pSHP2::ARF2 RNAi* and wild-type plants.
(a) Primer orientation within the *ARF2* gene. Grey boxes represent exons within the *ARF2* coding region, with exon numbers given below. (b) RT-PCR following RNA extraction from 3 DAP siliques of three *pSHP2::ARF2 RNAi* lines and Col-0.

4.3 Discussion

Arabidopsis arf2 mutants provide an opportunity to test whether integument-led seed growth can enhance seed yield. However *arf2*, like many other single gene mutations that alter seed size, also causes pleiotropic effects throughout development such as poor self-fertility, which may be undesirable from an agronomic perspective. This chapter describes three strategies to circumvent the problem of low fertility in *arf2* mutants to enable meaningful yield assessment.

4.3.1 arf2-9 heterozygotes do not have enhanced yield

Homozygous *arf2-9* mutants produce enlarged seeds, but the total yield per plant is severely reduced because of decreased seed production. Low seed set in the mutants is associated with the failure of flowers to open and self-pollinate as a result of the over-elongation of the sepals and gynoecia. The first strategy attempted to overcome the problem of reduced fertility by using *arf2-9* heterozygotes, which have normal flowers but retain some aspects of the homozygous mutant phenotype. *arf2-9* heterozygotes did not show greater total seed yield per plant than the wild-type in unrestricted pollinations, yet produced significantly larger seeds than the wild-type in restricted pollinations (Table 4.1). *arf2-9* heterozygotes and wild-type plants showed a similar dry mass of aerial organs (Figure 4.8b), ruling out increased biomass as an explanation. However, some seed coat layers of heterozygotes contained more cells than the wild-type at some time points (Figure 4.5), and this could provide a mechanism for increased seed size. It is possible that the effect of having only one wild-type copy of the *ARF2* gene is amplified by the excess of maternal resources available to each seed when seed set is restricted.

A previous study showed that *ap2-10* heterozygotes produced a higher seed yield than wild-type plants (Jofuku *et al*., 2005), therefore the seed size and yield of *ap2-7* and *ap2-11* heterozygotes was tested alongside *arf2-9* heterozygotes in order to compare results. Following restricted pollination, both *ap2-7* and *ap2-11* heterozygotes produced significantly smaller seeds than the wild-type, which may be a result of the large number of seeds per silique (Table 1). In contrast with the previous study in which *ap2-10* heterozygotes showed a 35% increased total seed yield compared with wild-type (Jofuku *et al*., 2005), both *ap2-7* and *ap2-11* heterozygotes did not show a greater total seed yield (Table 1). Different growth conditions could cause the variation in results. Alternatively, the use of different *ap2* mutant alleles may explain the discrepancy; but this is unlikely as *ap2-7*, *ap2-10* and *ap2-11* have all been reported to have similar phenotypes (Jofuku *et al*., 2005; Ohto *et al*., 2005). However, *ap2-7* and *ap2-11* mutant alleles are in the Col-0

background, whereas *ap2-10* is in C24; therefore, it is possible that the ecotype could account for a difference in behaviour.

4.3.2 Restoration of flower opening in arf2-9 improves fertility and seed yield

To repair the poor fertility of *arf2-9* mutants and subsequently carry out meaningful yield analysis, the second strategy involved re-instating *ARF2* expression in *arf2-9* mutants under the control of the *AP1* promoter. It is reasonable to assume that *AP1*-driven *ARF2* expression would restore *ARF2* function to the sepals and petals of *arf2-9* mutants, allowing flower opening; however as the *AP1* promoter does not cause gene expression elsewhere in the plant (Mandel *et al*., 1992), all other aspects of the *arf2* mutant phenotype should remain. It has previously been shown in *Arabidopsis* that seed size QTLs co-localise with other life history traits (Alonso-Blanco *et al*., 1999), and therefore it is possible that the production of large seeds in *arf2* mutants requires factors in addition to integument over-proliferation, which would be desirable to retain in plants with restored fertility. For instance, *arf2* mutants have enlarged rosette leaves which may therefore have increased photosynthetic capacity. Rosette leaves are the major source tissue in *Arabidopsis* (Robinson and Hill, 1999). Enlarged sepals and gynoecia may also be required for increased seed size in *arf2* mutants, as flowers and elongating siliques are key sink organs (Robinson and Hill, 1999), drawing resources from vegetative organs. Furthermore, the reproductive parts themselves can photosynthesize (Bazzaz *et al*., 1979). Silique wall photosynthesis is the main source of assimilates after leaf senescence occurs and, in *Brassica*, may contribute up to 60% of final plant dry matter (King *et al*., 1997). A study of *abscisic acid insensitive 3* (*abi3*) mutants in *Arabidopsis* showed that extra resources were available for seed production as a result of delayed senescence altering resource partitioning (Robinson and Hill, 1999). Hence, the delayed senescence of leaves and floral organs of *arf2* mutants (Ellis *et al*., 2005) may also contribute to increased seed size.

Flower opening requires the forced separation of the sepals by expanding petals (Van Doorn and Van Meeteren, 2003). Plants containing the *pAP1::ARF2* transgene had wild-type-like sepals and petals and the flowers were able to open. Fertility was substantially improved relative to *arf2-9*, with seed set increased from 20.5% to 78.2%. Despite comparatively normal sepals, petals and stamens in flowers of *pAP1::ARF2* plants, the gynoecia remained over-elongated. Increased fertility in *pAP1::ARF2* plants could therefore be caused by two factors: the reduced sepal length, which allows petal-driven flower opening, and over-elongation of the stamens, which brings pollen to the stigma on the enlarged gynoecium. Crucially, seeds in repaired plants also retained the oversized

integument and seed phenotype. Therefore, increased cell division in the gynoecium and integuments does not depend on general floral overgrowth.

In both restricted and unrestricted pollinations, *pAP1::ARF2* seeds were larger than those of the wild-type. In restricted pollination, seed set was held constant, and hence the observed difference between *pAP1::ARF2* and wild-type seed size was a result of altered seed development rather than differences in fertility. In unrestricted pollination, *pAP1::ARF2* seeds were also larger than those of the wild-type despite much improved fertility. However, fertility restoration was incomplete, as fewer seeds were produced per silique than in the wild-type. Although the stamen to gynoecium ratio indicated that the anthers were brought into the vicinity of the stigma, it appears that insufficient pollen was deposited on to the stigma to fertilize all ovules. This highlights the delicate balance between the lengths of all the floral organs required for a plant to be fully fertile, especially in the absence of insect pollinators. Although the discrepancy between stamen and gynoecium length causes a loss of fertility in *Arabidopsis*, seed set should not be affected in *B. napus* which uses both insects and wind to achieve successful pollination. Experiments are now underway to test the impact of integument-led seed growth on yield in this important crop species (Collaboration between: Rod Scott, University of Bath; Graham King and Smita Kurup, Rothamsted Research).

As *arf2-9* seeds were altered in development and size, it was possible that the oil content and composition may have been modified by the mutation, with potentially important consequences for target crops, such as *B. napus*. Although the lipid analysis showed that *arf2-9* seeds contained more oil than wild-type seeds, this was a result of the greater seed size and not a higher percentage of fatty acids per microgram of seed (Table 4.3). This conclusion has been corroborated by recent quantitative genetic studies on several hundred accessions of *Arabidopsis* (O'Neill *et al*., 2003) and *Brassica* (Barker *et al*., 2007), where seed size is weakly correlated with oil yield, but no common loci controlling both seed size and oil yield have been identified. In addition, no dramatic changes were found to the oil profile of *arf2-9* seeds (Figure 4.7). These data may prove to be useful if the technology is transferred to oilseed crops, such as *B. napus*, in which the maintenance of consistent fatty acid ratios independent of oil yield may be important.

The HI of *arf2-9* mutants was dramatically improved from 0.01 to 0.195 by the introduction of the *pAP1::ARF2* transgene. Although enlarged leaves, delayed flowering and increased stem thickness were unaffected by fertility restoration in *pAP1::ARF2* plants (Table 4.2), there was a decrease in the dry mass of aerial organs (Figure 4.8b). Inflorescence stems were shorter in *pAP1::ARF2* than in *arf2-9* plants, and it is also possible that there were effects on characters not assayed in this study, such as leaf size or

the composition of cells. The carbon cost of seed production is considerable (Munier-Jolain and Salon, 2005), especially in oilseeds such as *Arabidopsis* and *B. napus*. Oil has a production value index (PVI) of 3.03, more than double the carbon cost associated with the synthesis of cell wall and starch, which both have a PVI of 1.21 (Penning De Vries *et al*., 1974; Vertregt and Penning De Vries, 1987). The total metabolic cost of producing seeds includes not only the cost of the synthesis of storage products (proteins, lipids and starch), but also the cost of nutrient translocation from source to sink organs. Hence, the proportion of energy allocated to reproduction in the plant is greater than is assumed by simply assessing the ratio of seed weight to total plant biomass. As a result of the low fertility of *arf2-9* homozygous mutants, few resources are allocated to seed production, and therefore extra energy should be available for further vegetative development. Consistent with this, *ap2* homozygous mutants had a significantly increased dry mass of aerial organs compared to wild-type even though these mutants do not show increased cell division or organ overgrowth.

4.3.3 INO-driven suppression of ARF2 does not significantly impact seed yield

The introduction of the *pINO::ARF2 RNAi* transgene into plants did not cause an increase in seed weight in unrestricted pollinations or an increase in the total seed yield (Table 4.5). The result was the same regardless of whether the plants contained single or multiple copies of the transgene, suggesting that the level of *ARF2* transcript has not been sufficiently reduced in the integuments of the transgenic plants to enhance seed size. Previously, the same *ARF2* RNAi fragment has been placed under the control of the constitutive CaMV 35S promoter, yielding transgenic plants nearly indistinguishable from *arf2-9* mutants (Schruff, 2006). Crucially, the large-seed phenotype seen in *arf2* mutants was reproduced using RNAi technology. Despite a visible phenotype in *p35S::ARF2 RNAi* plants, a decrease in *ARF2* expression was not detectable by semi-quantitative RT-PCR. *ARF2* is expressed throughout young ovules, but in mature ovules expression remains high in the integuments, a small group of nucellar cells and in the funiculus (Schruff, 2006). No expression was detected in the seed coat following fertilization. It is difficult to determine the crucial point during this expression profile at which *ARF2* exerts its influence over cell division in wild-type integuments. In contrast to the relatively broad expression of *ARF2* in the developing ovule, the *INO* promoter has limited expression within the ovule primordium and in the elongating tissue of the outer integument on the abaxial side of the chalaza (Villanueva *et al*., 1999). Therefore, despite *INO* having a known role in early ovule development, *ARF2* expression may not have been altered sufficiently to disrupt integument development and increase seed size.

Although no difference in the seed size of *pINO::ARF2 RNAi* plants was detected in unrestricted pollinations, significantly larger seeds were produced in restricted pollinations (Table 4.5). Fertility, as judged by seed number per silique, was uncompromised in the transgenic plants, thus the observed increase in seed size did not result from reduced seed number. It is possible that any effect exerted by the *pINO::ARF2 RNAi* transgene on seed development may be amplified in restricted pollinations.

The HI of plants containing the *pINO::ARF2 RNAi* transgene was lower than that of the wild-type, and this was primarily due to an increase in the dry mass of aerial organs (Table 4.5). It is unknown why *pINO::ARF2 RNAi* plants should have a higher dry mass of aerial organs than wild-type. One possible explanation is that the *INO* promoter is expressed during vegetative development as well as within the ovule integuments; yet this is unlikely as plants homozygous for *ino* mutations are vegetatively normal, although they do exhibit general effects associated with reduced fertility (Baker *et al*., 1997). However, in situ hybridisation using an *INO* antisense probe has revealed that *INO* mRNA is not only found in ovules, but may also be present in globular embryos (Villanueva *et al*., 1999).

4.3.4 SHP2-driven ARF2 suppression increases seed weight but reduces fertility

In both restricted and unrestricted pollinations, *pSHP2::ARF2 RNAi* plants produced seeds that were larger than those of the wild-type (Figure 4.12). However, in restricted pollination the flowers were not manually pollinated and seed set was not complete; hence the increase in seed weight may be caused in part by the decrease in seed number. As the *SHP2* promoter directs gene expression in the carpel, as well as in ovules and developing seeds (Savidge *et al*., 1995; Dinneny *et al*., 2005), a reduction in *ARF2* mRNA in the gynoecium could lead to the over-elongation of this floral organ. Fertility may be compromised as the stamens cannot unload sufficient pollen on to the stigma of the over-elongated gynoecium to fertilise every available ovule. Although, *ARF2* suppression throughout the carpel is associated with negative effects on fertility in *Arabidopsis*, the overgrowth of the gynoecium would not be a problem in insect or wind-pollinated plants, such as *B. napus*. The enlarged gynoecium may be necessary for increased seed size as the enlarged silique may provide the space for further seed growth. Furthermore, silique wall photosynthesis may be increased, providing more assimilates for seed filling (Bazzaz *et al*., 1979; King *et al*., 1997).

Three independent *pSHP2::ARF2 RNAi* lines were analysed for seed weight and fertility and varied results were obtained. *pSHP2::ARF2 RNAi* -5 and *pSHP2::ARF RNAi* -26 showed high seed weights and low fertility whereas *pSHP2::ARF2 RNAi* -22 showed a less dramatic phenotype, with seed weight and seed set similar to that of wild-type plants.

Consistent with this, the level of *ARF2* transcripts, as determined by semi-quantitative RT-PCR, was visibly reduced in lines 5 and 26 compared to wild-type but not in line 22 (Figure 4.13). Varying levels of gene suppression and phenotype severity between transgenic lines has been reported previously when using RNAi (Chuang and Meyerowitz, 2000). RNAi has been described as a knockdown, rather than a knockout strategy (Kusaba, 2004), so it is not surprising that *ARF2* transcripts were detected in the three *pSHP2::ARF2 RNAi* lines. In addition, RNA was extracted from 3 DAP siliques, in which *SHP2*-driven expression is limited to the valve margins and developing seeds (Savidge *et al*., 1995; Dinneny *et al*., 2005). Therefore, *ARF2* expression would not be suppressed in all regions of the tissue assayed.

4.3.5 Further work

To determine whether targeted *ARF2* suppression, either directed by the *INO* or the *SHP2* promoter, modifies seed development and size through increased cell division, it is necessary to carry out more detailed phenotypic analyses of the transgenic plants. The number of cells in the integument layers of both *pINO::ARF2 RNAi* and *pSHP2::ARF2 RNAi* seeds should be counted and compared to wild-type. In addition, the floral organs of *pSHP2::ARF2 RNAi* transgenic plants should be measured to establish whether reduced fertility is associated with over-elongation of the gynoecium. The analysis of *pSHP2::ARF2 RNAi* seed weight following restricted pollinations should be repeated and full seed set ensured by manual pollination.

The work presented here uses *Arabidopsis* and the *arf2* mutant as a model. The aim of this project was develop technology that could be transferred to economically important crop species, such as *B. napus*. The results from both the *pAP1::ARF2* and *pSHP2::ARF2 RNAi* show that biotechnological approaches to increase seed size by modification of the ovule integuments has the potential to increase seed yield. Further yield trials in *B. napus* are necessary to determine whether yield enhancement can be achieved, without the complication of reduced fertility.

4.3.6 Summary

Homozygous *arf2* mutants have enlarged seeds but poor fertility. In order to use the *arf2* mutant as a model to test whether integument-led seed growth can enhance yield, three strategies were employed to overcome the confounding effect of reduced fertility.

1. Semi-dominant heterozygotes. *arf2* heterozygotes produced significantly larger seeds than wild-type in restricted pollinations, but in unrestricted pollinations seed size and total seed yield were not increased.

2. Restoration of flower opening in arf2-9 mutants. Targeted restoration of the *arf2* mutant phenotype using the *pAP1::ARF2* construct enabled flower opening and self-pollination, considerably improving *arf2* mutant fertility whilst maintaining increased seed size. Compared to *arf2* mutants, *pAP1::ARF2* plants had a dramatically greater seed yield and HI. *arf2* and *pAP1::ARF2* seeds contained more oil than wild-type seeds due to their increased size; however no dramatic changes to the fatty acid profile were observed.

3. Increasing seed size through targeted suppression of ARF2. *SHP2*-driven suppression of *ARF2* in wild-type plants increased seed size in both restricted and unrestricted pollinations. This increase in seed weight may in part be caused by a reduction in fertility, likely a result of an over-elongated gynoecium. Targeted suppression of *ARF2* in wild-type plants using the *INO* promoter did not significantly alter mean seed weight or total seed yield.

These experiments highlight the importance of the integuments/seed coat in determining seed yield. In addition, *ARF2* has been shown to be a useful target gene for modifying integument development as well as for increasing plant growth and total biomass.

5. THE SEARCH FOR MORE REGULATORS OF SEED SIZE AND SEED SHAPE

5.1 Introduction

5.1.1 Suppressor/enhancer screens

Forward genetic screens enable the identification of genes involved in many aspects of plant development. Particularly powerful are suppressor/enhancer screens which involve the induction of second-site mutations in plants which already carry a known mutation. This type of screen can uncover more information about a gene or mutation, and identify other interacting components in a pathway (Page and Grossniklaus, 2002). In addition, redundant genes can be identified. Mutations in redundant genes possess silent phenotypes as other genes in the same or a parallel pathway can perform the same function. However, double mutants can reveal the role of these genes in a specific process. For example, second-site mutations were induced in the *Arabidopsis* mutant *crabs claw* (*crc*) which has defects in carpel development (Eshed *et al*., 1999). This led to the discovery that *gymnos* (*gym*)/*pickle* (*pkl*) and *kanadi1* (*kan1*) enhance the *crc* phenotype causing ectopic ovule formation on the outside of the carpels. In an extension of this study, a search was conducted for enhancers of the *pkl-12 kan1-2* double mutant and subsequently the novel locus *KAN2* was identified (Eshed *et al*., 2001). Both *kan1* and *kan2* single mutants display little or no effect on plant morphology, yet the double mutant shows an almost complete loss of carpel polarity. The use of suppressor/enhancer screens has therefore established a better understanding of how carpel polarity is determined in *Arabidopsis*.

This chapter describes a suppressor/enhancer screen designed to gain further insight into seed development and seed size regulation in *Arabidopsis*. Second-site mutations were induced in the *arf2* mutant, which produces large, pointed seeds associated with the presence of extra cells in the ovule integuments (Schruff *et al*., 2006). Mutant *arf2* plants also exhibit large leaves, thick and twisted inflorescence stems, and low fertility associated with over-elongation of the sepals and gynoecia which prevent flower opening and self-pollination (Okushima *et al*., 2005; Schruff *et al*., 2006). Enlarged organ size due to over-proliferation in *arf2* mutants suggests that the wild-type function of the ARF2 protein is to repress cell division. However, cell expansion is also affected by *ARF2* (Li *et al*., 2004; Schruff *et al*., 2006). In addition, *arf2* mutants are delayed in the onset of flowering, leaf senescence, floral organ abscission and silique ripening (Ellis *et al.*, 2005).

The screen was expected to identify two classes of second-site mutants:

1. Suppressors - Candidate mutants that appeared wild-type-like despite containing a mutation in the *arf2* gene.
2. Enhancers - Candidate mutants with exaggerated *arf2* phenotypes.

This screen could potentially identify genes encoding proteins that: interact directly with ARF2; act in the same pathway as ARF2; or are involved in similar processes, such as cell division, cell expansion and organ size control, but act independently of ARF2.

5.1.2 The ARF2 signalling network

ARF2 is a member of a family of transcription factors that regulate gene expression in response to the plant hormone auxin (Guilfoyle and Hagen, 2007). Auxin has been implicated in virtually every aspect of growth and development; however despite the diversity of cellular responses triggered by auxin, most are controlled by a relatively short pathway that alters gene expression (Vanneste and Friml, 2009). The main naturally-occurring auxin present in plant cells is indole-3-acetic acid (IAA). Auxin is perceived by a small family of F-box proteins including TRANSPORT INHIBITOR RESPONSE 1 (TIR1) (Dharmasiri *et al*., 2005; Kepinski and Leyser, 2005). Two classes of transcriptional regulators mediate gene expression; there are 29 genes encoding Aux/IAAs and 23 genes encoding ARFs (Guilfoyle and Hagen, 2007). Aux/IAAs are short lived, nuclear proteins which cannot bind DNA directly but modulate transcription by dimerising through a conserved domain with ARF proteins (Ulmasov *et al*., 1999). ARF proteins bind DNA at particular auxin-responsive elements (AuxREs) in the promoters of auxin responsive genes (Guilfoyle *et al*., 1998).

It is believed that distinct pairs of ARF and Aux/IAA proteins convert a generic auxin signal into a specific response (Weijers *et al*., 2005). Some auxin responsive genes are up-regulated and some down-regulated via this system, depending on whether the ARF protein acts as an activator or a repressor. This is determined by the amino acid composition of the middle region of the protein (Tiwari *et al*., 2003). ARF2 is known to function as a repressor. At low concentrations of auxin, Aux/IAAs bind activating ARFs and thereby repress transcription of auxin responsive genes. When present at a high concentration, auxin acts as ‘molecular glue’ enhancing the binding of TIR1 to the Aux/IAA, which is targeted for ubiquitin-mediated degradation by SCF^{TIR1} E3 ligase (Tan *et al*., 2007). Aux/IAA degradation permits activating ARFs to promote the expression of auxin responsive genes. In contrast, it is proposed that repressing ARFs like ARF2

compete with activating ARFs in a concentration-dependent manner to bind AuxREs (Guilfoyle and Hagen, 2007).

ARF2 has been shown to interact with other auxin signalling components. For example, ARF2 can bind directly to ARF1 (Ulmasov *et al*., 1997) which is also a transcriptional repressor (Tiwari *et al*., 2003). The *arf1* mutant alone shows no obvious phenotype but *arf1 arf2* double mutants displayed enhanced effects of the *arf2* mutation on apical hook formation (Li *et al*., 2004), flowering time and senescence (Ellis *et al*., 2005). Delayed senescence in *arf2* mutants was also enhanced by mutations in *ARF7* and *ARF19*, although these two transcription factors did not affect senescence on their own (Ellis *et al*., 2005). Despite these connections, it has been proposed that *ARF2* does not fit the canonical auxin response model as global expression of auxin-regulated genes was not altered in *arf2* mutants (Okushima *et al*., 2005; Ellis *et al*., 2005). Supporting this, there is evidence suggesting that *ARF2* affects the expression of other types of genes. For example, in *arf2* mutants expression of the *ACS* gene family (involved in ethylene biosynthesis) is inhibited in flowers (Okushima *et al*., 2005) and *SENESCENCE ASSOCIATED GENE 12* (*SAG12*) is repressed in senescing leaves (Ellis *et al*., 2005). Additionally, the expression of *ANT* and *CYCD3;1* genes that promote cell division is prolonged in *arf2* mutants (Schruff *et al*., 2006). ARF2 has also been shown to be bound by BRASSINOSTEROID INSENSITIVE 2 (BIN2), a brassinosteroid-regulated kinase which phosphorylates ARF2 reducing its ability to bind DNA and repress transcription (Vert *et al*., 2008).

In summary, ARF2 has been shown to interact with a diverse group of proteins and consequently is believed to play a role in several signalling pathways. Therefore, a large number of potential targets which could either suppress or enhance the *arf2* mutant phenotype should exist. This screen provides an opportunity to identify these target genes, particularly those which have a role in seed development.

5.1.3 Ethylmethane sulphonate (EMS) mutagenesis

In this study, the chemical mutagen EMS was used to artificially induce heritable mutations in the *Arabidopsis arf2* mutant. EMS is an alkylating agent, modifying guanine (G) residues to O^6-ethylguanine which base pairs with thymine (T) rather than cytosine (C) (Greene *et al*., 2003). If this modification is not repaired, DNA replication will replace the G/C base pair with an A/T. EMS is a popular choice of mutagen as doses that induce frequent single base changes permit high rates of plant survival (Malmberg, 2004). In *Arabidopsis*, the most practical method to induce mutations is through seed mutagenesis with the aim of targeting the diploid cells of the embryo (Page and Grossniklaus, 2002). Since the embryo is not a single cell, it will be chimeric for wild-type cells and mutant

cells. The resulting M_1 plant will therefore contain sectors, some of which will be heterozygous for a particular mutation. For the mutation to be inherited by the next generation, it must be present in one of the cells destined to become a gamete. Following self-fertilisation, the homozygous mutants will segregate in the M_2 generation. In *Arabidopsis*, recessive mutants segregate at a 7:1 ratio, suggesting that there are two genetically effective cells in the embryo that ultimately give rise to gametes.

In this suppressor/enhancer screen, 5,000 *arf2* seeds were mutagenised. Although great care was taken to ensure that these seeds were all homozygous for the mutant *arf2* allele, the presence of *arf2* heterozygote or wild-type seeds could not be ruled out. As these contaminating seeds would produce plants that could be confused with interesting suppressors of *arf2*, it was important to easily distinguish false positives. Therefore, a T-DNA insertion line designated Salk_108995 (Alonso *et al*., 2003), also referred to as *arf2-8*, was used as the genetic background in this screen. The *arf2-8* mutant allele carries a T-DNA insertion in the twelfth exon at nucleotide position 3,097 (Okushima *et al*., 2005), and since *ARF2* mRNA was not detected in *arf2-8* plants they are considered null mutants (Ellis *et al*., 2005). Homozygous *arf2-8* mutants show the same phenotype as *arf2-9*/*mnt* mutants (Schruff *et al*., 2006), but have the advantage that the mutant allele can be tracked using PCR (Figure 5.1). The 'lba1' primer of the T-DNA insertion was used together with a primer ('0.05F') designed specifically for the *ARF2* gene to check for presence of the T-DNA insertion. A second PCR, using two primers ('0.05F' and '0.05R') within the *ARF2* gene situated on either side of the T-DNA insertion, indicated whether a wild-type copy of the gene was present (primers and PCR conditions are detailed in Table 2.1 and Table 2.2). PCR genotyping therefore allowed false positives to be eliminated as *arf2* heterozygotes produced positive results for both PCRs and wild-type plants tested negative for the presence of the T-DNA insertion.

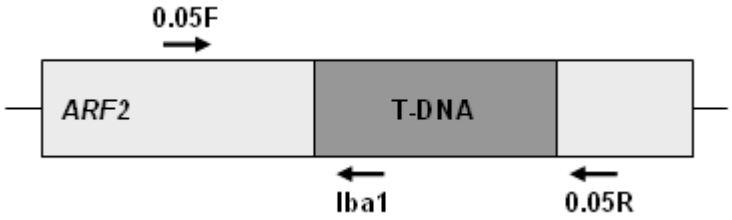

Figure 5.1: Genotyping *arf2-8*/Salk_108995 T-DNA insertion mutants.
Location of primer binding sites in *arf2-8* (primer design by M.C. Schruff, University of Bath).

5.1.4 Gene discovery

Once suppressor or enhancer candidates of the *arf2* phenotype have been identified, high resolution gene mapping provides a powerful tool to identify the mutated genes responsible. Until recently, map-based cloning was considered an extremely time-consuming and expensive process. However, the publication of the entire sequence of the *Arabidopsis* genome along with annotated molecular markers has facilitated the mapping process remarkably (Lukowitz *et al*., 2000). A typical fine mapping experiment now takes little more than nine months, assuming no complications are encountered.

5.2 Results

5.2.1 EMS mutagenesis and screening

A pilot test was carried out to find the optimum dose of the mutagen as *arf2-8* seeds may require a different EMS dose than wild-type due to their large size. The presence of chlorophyll-deficient sectors in the M_1 plants indicated successful mutagenesis (Figure 5.2a). M_1 plants treated with a high concentration of EMS were less vigorous than those treated with a low concentration showing the increasing severity of mutagenesis (Figure 5.2b-d). The optimum dose of EMS gives a high mutation level but does not cause sterility. This was determined by calculating the percentage of M_1 plants with siliques that carry roughly 25% embryo lethals, which should ideally lie between 10% and 30% (Luca Comai, personal communication, 2005). As M_1 plants are chimeric with both wild-type and mutant sectors, multiple siliques from different inflorescences were screened for aborted seeds on each plant. The aborted seeds represent the offspring homozygous for mutations in genes that lead to embryo lethality and due to the recessive nature of most mutations they first appear in the M_2 generation. The frequency of embryo lethal (*emb*) mutations is used to measure the success of mutagenesis as they are the most common heritable defect following *Arabidopsis* mutagenesis (Kyjovska *et al*., 2003) and are relatively easy to score. (Figure 5.2e). Results of the pilot test indicated that 0.2% treatment was the optimum dose as 33% of siliques contained embryo lethals (Figure 5.2f).

A large scale mutagenesis of 5,000 seeds treated with 0.2% EMS was then carried out. M_2 plants were scored for three aspects of the *arf2* phenotype: variations in seed size and shape, flower opening and fertility. 800 plants from each of five families were screened yielding twenty-one candidates with altered seed size and shape (Figure 5.3). These candidates could be categorised into three groups: seeds that appeared rounder than *arf2-8* seeds and resembled wild-type seeds (suppressors), seeds that appeared more pointed than *arf2-8* seeds (enhancers) and a third group containing those seeds which could not be classified as either suppressors or enhancers, but had altered seed pigmentation. The candidates in the latter group produced pale seeds and are likely to be *transparent testa* (*tt*) mutants, defective in genes required for various steps in flavonoid biosynthesis (Koorneef, 1990, Shirley *et al*., 1995) A few plants with open flowers and full fertility were identified but found to be false positives, as PCR genotyping revealed that they were not homozygous for the *arf2-8* allele.

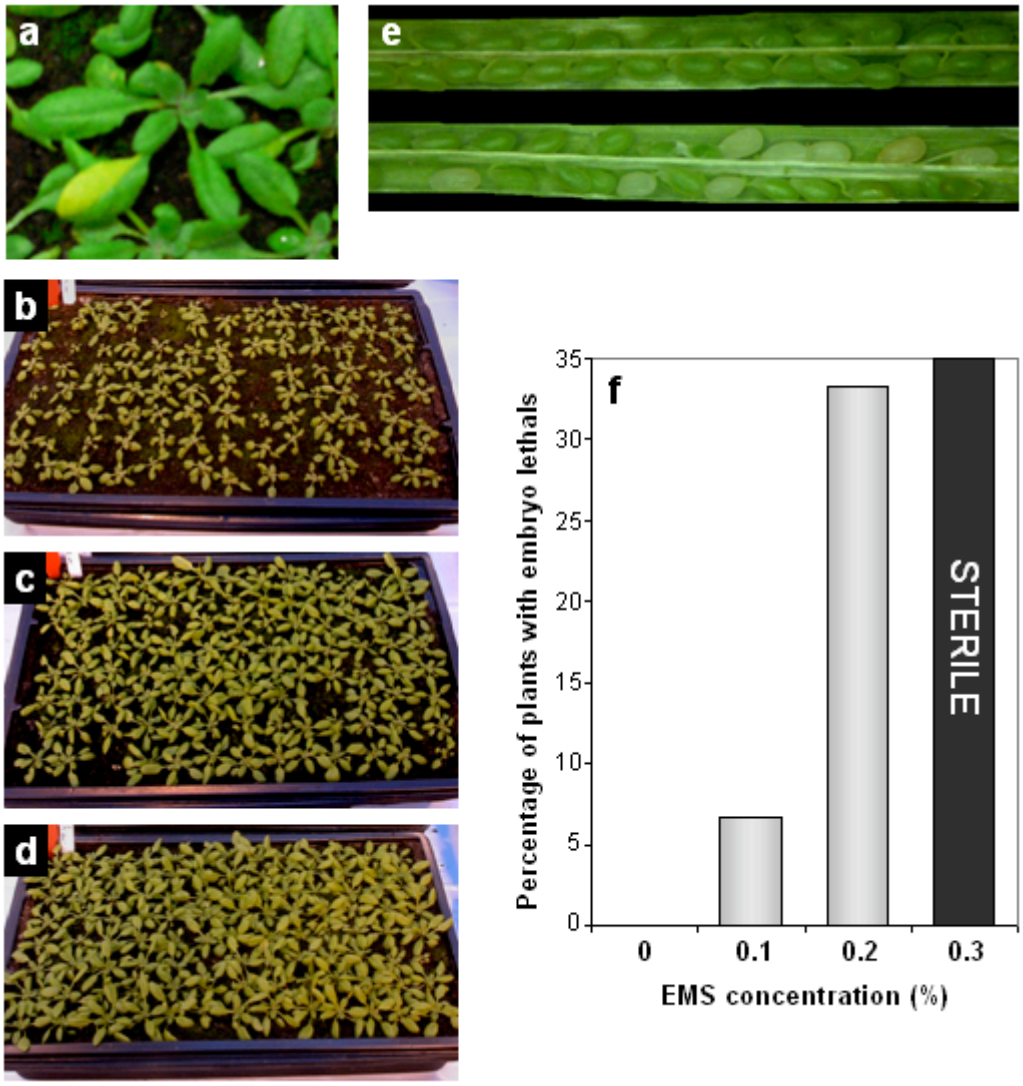

Figure 5.2: EMS mutagenesis

(a) Chlorophyll deficient sectors in M_1 plants. M_1 plants appeared healthier with decreasing EMS concentration from 0.3% (b), 0.2% (c) to 0.1% (d). (e) Siliques containing healthy (top) and aborted seeds/embryo lethals (bottom). (f) Graph showing the percentage of siliques carrying roughly 25% embryo lethals at varying EMS concentrations. n = 30.

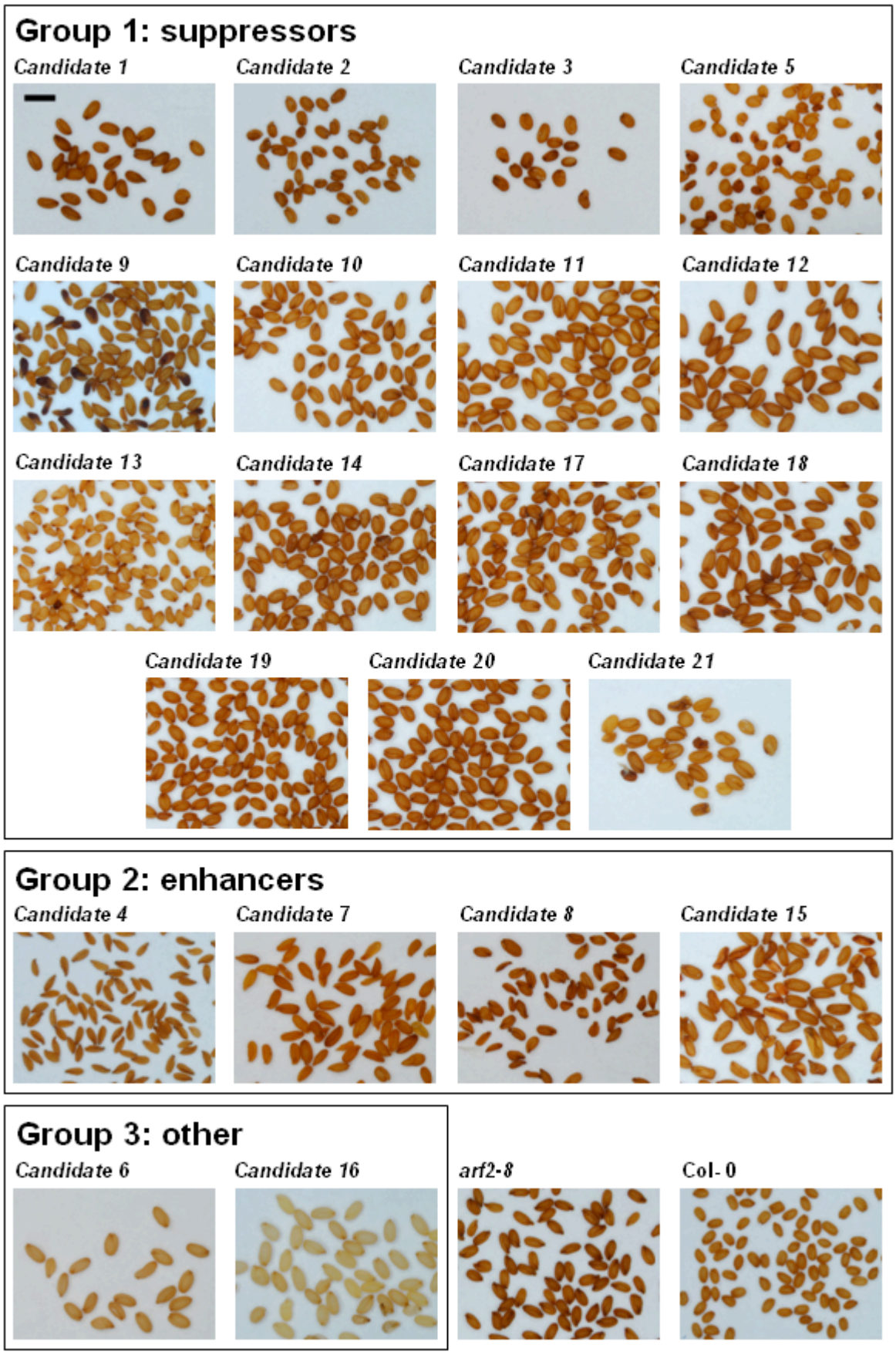

Figure 5.3: Seed phenotypes of EMS-induced candidate lines.

M_3 generation seeds were screened for altered seed shape and size compared to *arf2-8* mutant seeds and classified into three groups: suppressors of *arf2*, enhancers of *arf2* and those which do not fit into the other two groups. Bar = 1 mm.

From the twenty-one candidate lines selected on seed phenotype, a subset of lines was taken forward to vegetative phenotype analysis. The chosen lines had at least one of the following: a consistent seed shape (round or pointed); good fertility; and low levels of shrivelled or aborted seeds. At this stage all the candidates were PCR genotyped to confirm that they had remained homozygous for *arf2-8*. For each candidate line eight plants were assessed for vegetative phenotype (Figure 5.4). It is important to note that the candidate lines may contain more than one EMS-induced mutation. As most mutations are recessive, it is reasonable to assume that the mutations causing the seed phenotype are homozygous, and therefore should not segregate in the next generation. However, other mutations may segregate in the M_3 generation and therefore the plants from a single candidate line could display different vegetative morphologies. Indeed, *Candidates 7* and *15* contained plants with varied rosette phenotypes. Some candidate lines produced plants that were similar in appearance to *arf2-8*, such as *Candidate 9*, whereas as others displayed strikingly altered phenotypes, such as *Candidate 11*.

In order to choose a single candidate line for gene mapping the following questions were addressed:

1. Is the seed phenotype heritable? The phenotype of M_4 seeds was compared to seeds from the previous generation.
2. Is the mutation dominant or recessive? Candidate lines were backcrossed to *arf2-8* mutant plants and the F_2 generation screened for seed phenotype. If 25% of plants produced seeds that resembled the candidate phenotype the mutation was assumed to be recessive. These crosses also revealed whether any vegetative phenotype was linked to the seed phenotype. Co-segregation would suggest that a single gene mutation was causing pleiotropic affects throughout plant development. Further backcrosses into *arf2-8* were also carried out to remove any other EMS-induced mutations that were not associated with the altered seed phenotype.
3. Is the seed phenotype visible in a wild-type genetic background? Candidate lines were crossed with Col-0 to remove the *arf2-8* mutation. The seed phenotype of F_2 plants carrying two wild-type copies of *ARF2* (as determined by PCR genotyping), was assessed for variations from wild-type.

An 'ideal' candidate would possess a heritable seed phenotype and be associated with a single recessive mutation. Although it would be beneficial for the seed phenotype to be visible in a Col-0 background for mapping purposes, a seed phenotype present only in the *arf2-8* mutant background is more likely to result in the identification of a previously uncharacterised gene.

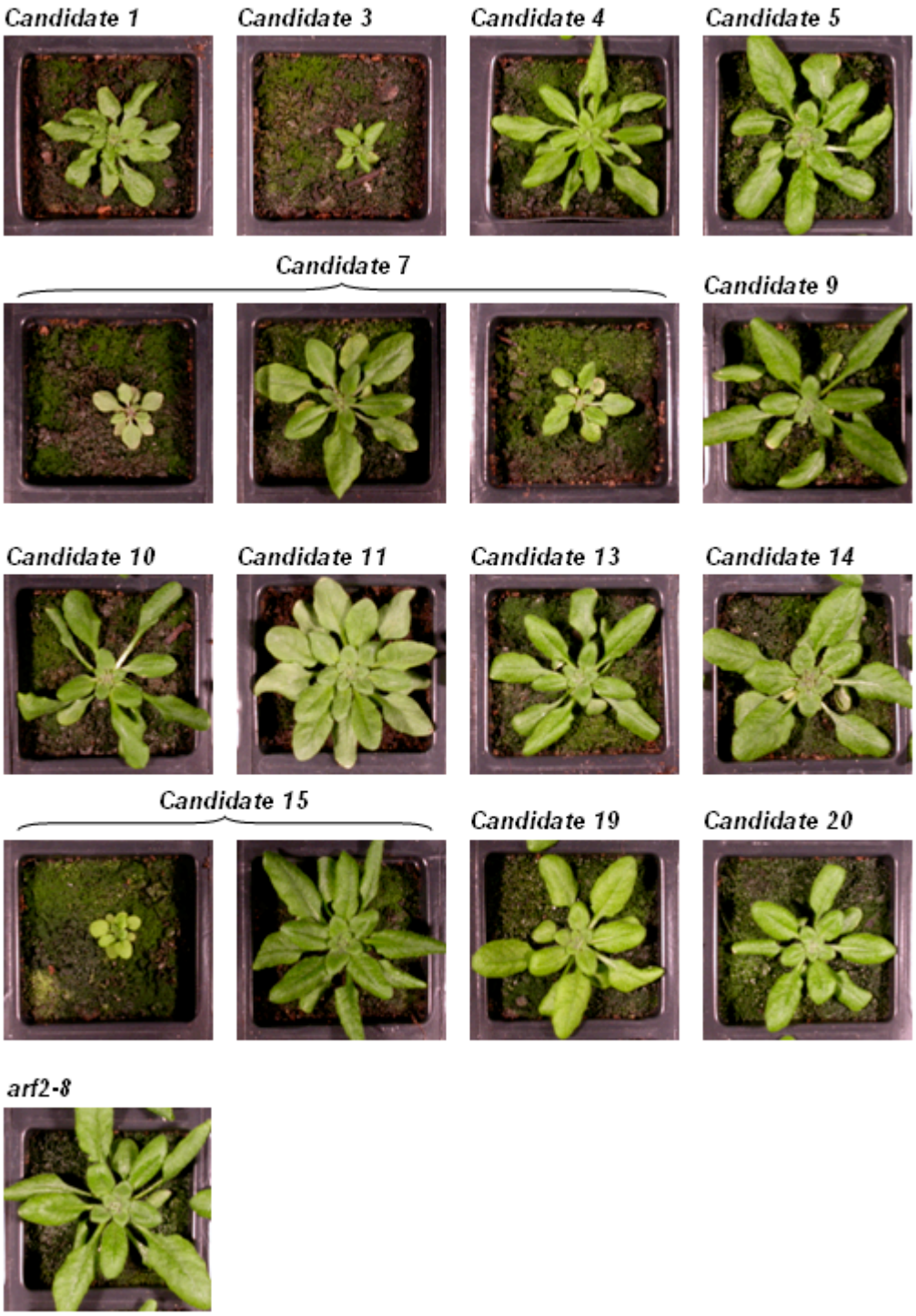

Figure 5.4: Rosette morphology of EMS-induced candidate lines.

M_3 generation plants at 30 days after germination. For candidate lines in which all plants showed similar morphology only one photograph is given. Where a variation within a line was found, the photographs represent the range of rosette morphologies present.

5.2.2 Candidate 3: a suppressor of the arf2 seed phenotype

Candidate 3 was considered a suppressor of *arf2* as plants in this line produced rounder seeds than the *arf2-8* mutant (Figure 5.3). Mature seeds of *Candidate 3* in both the M_3 and M_4 generations were round; therefore an EMS-induced mutation within these plants had caused a heritable alteration to the seed phenotype. Further analysis revealed that roundness in *Candidate 3* seeds was associated with abnormal development of the ii1' seed coat layer, which was often two-cells thick (Figure 5.5). Additionally, the adaxial ridge, which is enlarged in *arf2-8* seeds, was greatly reduced in *Candidate 3* seeds. The M_3 and M_4 plants all showed the same vegetative phenotype, characterised by only a few small rosette leaves (Figure 5.4). Since many organs, including rosette leaves, are enlarged in *arf2* mutants, the vegetative and seed phenotype of *arf2-8* are both suppressed in *Candidate 3* plants.

A further phenotypic alteration observed in *Candidate 3* plants compared to *arf2-8* mutants, was abnormal carpel development (data not shown). Wild-type *Arabidopsis* gynoecia are composed of an apical stigma and two fused carpels containing the developing ovules (Ferrándiz *et al*., 1999). In *Candidate 3*, developing ovules were visible as carpels failed to fuse completely, and the stigma was enlarged. In addition, *Candidate 3* plants showed poor fertility, in part due to the *arf2-8* mutation which causes failure of early flowers to open and self-pollinate but also as a result of reduced pollen production (data not shown).

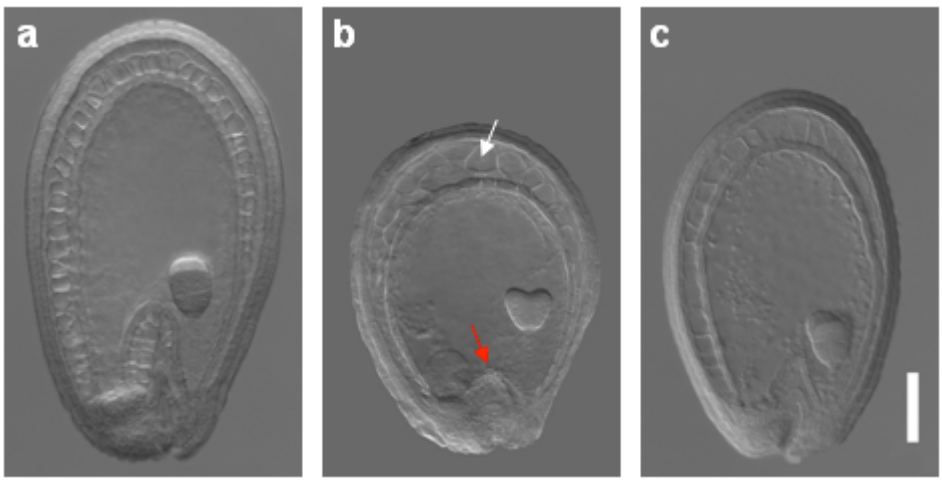

Figure 5.5: Seed phenotype of *Candidate 3* (*arf2-8* background).
Cleared seeds imaged with differential contrast optics from *arf2-8* (a), *Candidate 3* in the *arf2-8* background (b) and Col-0 (c). Photographs were taken on different occasions and are not of seeds of the same age; however they are of the same magnification and are given to demonstrate seed shape and seed coat morphology only. White arrow = ii1' seed coat layer. Red arrow = adaxial ridge. Bar = 100 μm.

The severe floral phenotype of *Candidate 3* plants reduced seed set to very low levels and hindered emasculation and pollination. *Candidate 3* was backcrossed to *arf2-8* mutant plants and a few F_1 seeds were obtained. Of the sixty F_2 plants that were grown, only one resembled the original *Candidate 3* mutant and produced round seeds. Some variation in rosette size was found, but only the one plant mentioned above was as small as in the original *Candidate 3* mutant. None of the other F_2 plants had round seeds, although many exhibited 25% seed abortion in their siliques. The embryo of these aborted seeds did not develop further than the eight-cell embryo stage. It is possible that the seed abortion was linked to the lack of *Candidate 3*-like plants in the F_2 population.

After crossing *Candidate 3* with Col-0, seventy-three F_2 plants were grown and genotyped for the presence of the *arf2-8* allele. Nineteen were found to be homozygous for the wild-type *ARF2* allele. None of these plants resembled *Candidate 3* or had round seeds. However, as Col-0 has seeds that are rounder than *arf2-8* mutants it would have been difficult to distinguish the *Candidate 3* seed phenotype in this background. Aborted seeds were also found in the siliques of the *Candidate 3* x Col-0 F_2 population. A single *arf2-8* heterozygote plant was however found that resembled *Candidate 3*. These findings suggested that the *Candidate 3* phenotype was not visible in a Col-0 background, although this was complicated due to the low frequency of the phenotype. Thus despite the interesting phenotype of *Candidate 3*, further work on this line was discontinued as the difficulties with crossing and complicated genetics did not facilitate gene mapping.

5.2.3 Candidate 4: an enhancer of the arf2 seed phenotype

The seeds of *Candidate 4* were narrower and more pointed than those of *arf2-8*, classifying this line as an enhancer of the *arf2* phenotype (Figure 5.3). *Candidate 4* mature seeds in the M_3 and M_4 generations showed an identical phenotype; indicating that a heritable change to the seed phenotype had been caused by EMS mutagenesis. There were no obvious alterations to the morphology of the seed coat to explain the difference in seed shape between *Candidate 4* and *arf2-8* (Figure 5.6). The adaxial ridge remained enlarged in *Candidate 4* mutants as in *arf2-8*. The M_3 and M_4 plants all showed the same vegetative phenotype, with slightly smaller, narrower and curled rosette leaves than *arf2-8* (Figure 5.4). As in *arf2-8* mutants, the inflorescence stems of *Candidate 4* plants were twisted however they did not appear thickened as in *arf2-8* alone. The floral phenotype of *Candidate 4* plants was similar to that of *arf2-8*. It was also observed that seed germination in *Candidate 4* was not always successful and that many seedlings had altered gravitropic responses (Figure 5.7).

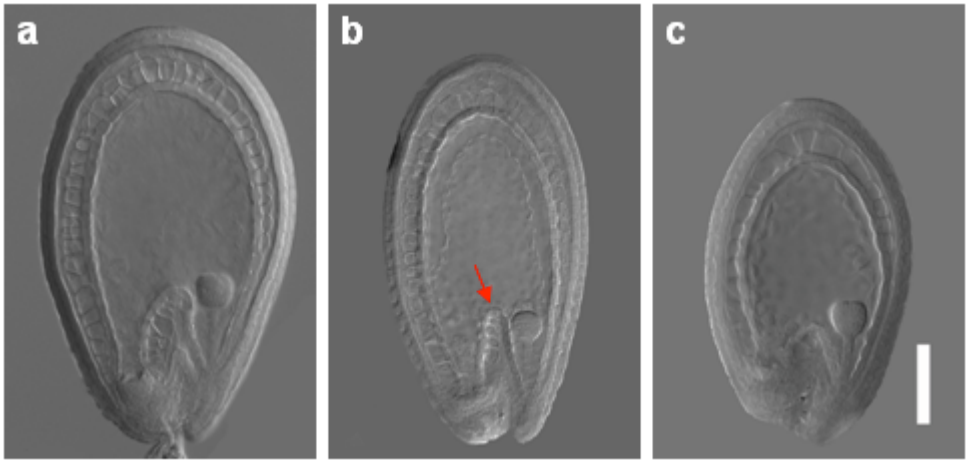

Figure 5.6: Seed phenotype of *Candidate 4* (*arf2-8* background).

Cleared seeds imaged with differential contrast optics from *arf2-8* (a), *Candidate 4* in the *arf2-8* background (b) and Col-0 (c). Photographs were taken on different occasions and are not of seeds of the same age; however they are of the same magnification and are given to demonstrate seed shape and seed coat morphology only. Red arrow = adaxial ridge. Bar = 100 µm.

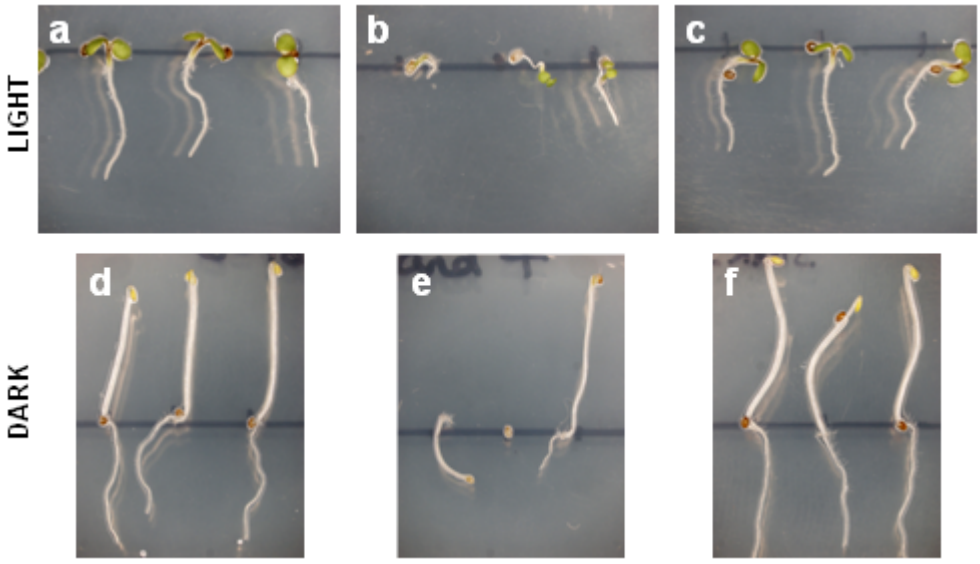

Figure 5.7: Phenotype of *Candidate 4* seedlings (*arf2-8* background).

Photographs of *arf2-8* (a and d), *Candidate 4* (b and e) and Col-0 (c and f) seedlings germinated in the light (a-c) and in the dark (d-f). Dark grown seedlings were exposed to light for 3 hrs to initiate germination. Plates were positioned 5º from vertical to encourage seedling growth along the surface rather than into the media.

Candidate 4 was backcrossed to *arf2-8* in order to determine the segregation ratio of the seed phenotype and remove other EMS-induced mutations that may be present. The *Candidate 4* seed phenotype was found to segregate as a single gene recessive mutation (8 *Candidate 4*-like and 22 *arf-8*-like plants were scored in the F_2 population; for a 3:1 segregation, $\chi^2 = 0.044$, P = 0.833).

The F_2 population resulting from the cross between *Candidate 4* and Col-0 was screened for both seed and vegetative phenotype. Twenty F_2 plants were genotyped for the *arf2-8* allele, five were wild-type, containing two copies of the *ARF2* allele. Two of these plants produced seeds that were wild-type-like in shape, yet small and pale. Eleven F_2 plants were found to be heterozygous for *arf2-8*, and three of these produced seeds that were wild-type-like in shape, yet small and pale (data not shown). The remaining four F_2 plants were *arf2-8* homozygotes and one plant produced seeds identical to those of the original *Candidate 4* mutant. This supported the previously stated assumption that *Candidate 4* is caused by a single gene recessive mutation (6 *Candidate 4*-like and 14 either wild-type-like or *arf2-8*-like were scored in the F_2 population, for a 3:1 segregation, $\chi^2 = 0.267$, P = 0.606). Thus, the *Candidate 4* phenotype was visible in a Col-0 background. The vegetative phenotype did not co-segregate with the seed phenotype, although it appeared more severe in the wild-type background (data not shown).

Candidate 4 seeds in the Col-0 background bear resemblance to those of *transparent testa glabra 2* (*ttg2*) seeds, which are small and pale (Johnson *et al*., 2002) due to defects in both integument and endosperm development (Garcia *et al*., 2005). A simple method to test whether *Candidate 4* is caused by a mutation in *TTG2* would be to perform an allelism test. If this is not the case, given that the *Candidate 4* seed phenotype is visible in both *arf2-8* and Col-0, gene mapping in either genetic background could be used to identify the gene of interest.

5.2.4 Candidate 14: a suppressor of the arf2 seed phenotype

Seeds of *Candidate 14* were rounder than *arf2-8* seeds hence this line was regarded as a suppressor of the *arf2* phenotype (Figure 5.3). The round seed phenotype was retained in subsequent generations indicating that an EMS-induced mutation had caused a heritable change in seed shape. Seed clearing revealed no obvious alterations to seed coat or endosperm development that could be responsible for the change in seed shape (Figure 5.8). In *Candidate 14* seeds, the adaxial ridge remained prominent as in *arf2-8* mutants. The M_3 and M_4 generation plants of *Candidate 14* showed the same vegetative phenotype, displaying leaves that were slightly broader than those of *arf2-8* (Figure 5.4). As in *arf2*-8, *Candidate 14* produced flowers that failed to open and self-pollinate due to over-elongation of the sepals and gynoecia.

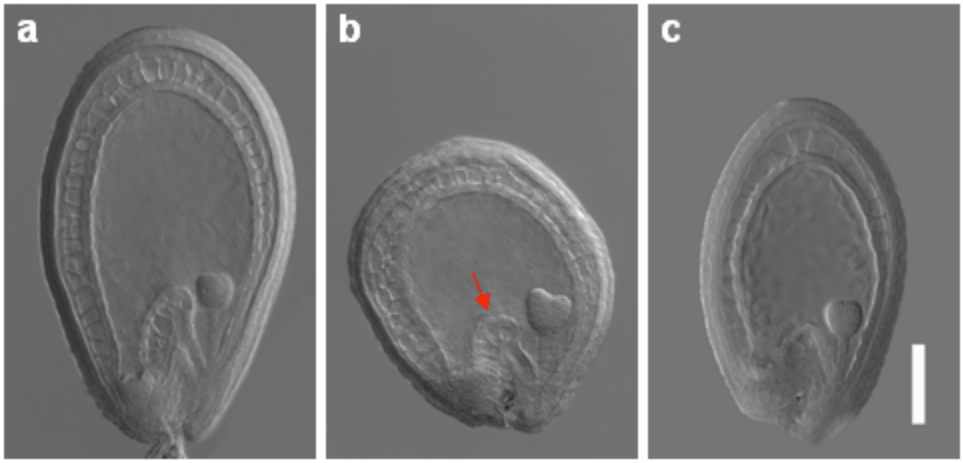

Figure 5.8: Seed phenotype of *Candidate 14* (*arf2-8* background).
Cleared seeds imaged with differential contrast optics from *arf2-8* (a), *Candidate 14* in the *arf2-8* background (b) and Col-0 (c). Photographs were taken on different occasions and are not of seeds of the same age; however they are of the same magnification and are given to demonstrate seed shape and seed coat morphology only. Red arrow = adaxial ridge. Bar = 100 μm.

Candidate 14 was backcrossed to *arf2-8* in order to determine the segregation ratio of the seed phenotype and remove other EMS-induced mutations that may be present. The *Candidate 14* seed phenotype was found to segregate as a single gene recessive mutation (9 plants with round seeds and 23 plants with pointed seeds were scored in the F_2 population; for a 3:1 segregation, $\chi^2 = 0.167$, $P = 0.683$). The vegetative phenotype did not co-segregate with the seed phenotype.

Candidate 14 was crossed with Col-0 to determine whether the seed phenotype was visible in the Col-0 background in addition to the *arf2-8* background. Thirty-two F2 plants were genotyped for presence of the *arf2-8* mutant allele. Of the eight plants found to contain two wild-type copies of the *ARF2* allele, two produced round seeds that were easily distinguished from wild-type seeds. Therefore, the *Candidate 14* seed phenotype was visible in the Col-0 background. Genotyping also revealed seventeen *arf2-8* heterozygotes (3 exhibited round seeds) and seven *arf2-8* homozygotes (2 exhibited round seeds). In total, seven in thirty-two plants were found to produce round seeds, supporting the assumption that the Candidate 14 seed phenotype was caused by a single gene recessive mutation (7 *Candidate 14*-like and 25 either wild-type-like or *arf2-8*-like were scored in the F_2 population; for a 3:1 segregation, $\chi^2 = 0.167$, $P = 0.683$).

Further analysis was subsequently carried out on the *Candidate 14* seed phenotype in the Col-0 background. The round seed phenotype of *Candidate 14* was easily distinguishable from Col-0 after manual silique opening (Figure 5.9a,b), and it was observed that *Candidate 14* siliques contained a large number of unfertilised ovules. Seed clearing revealed the presence of abnormal cells in the outer layer of the seed coat that protruded from the seed surface (Figure 5.9c). In addition, a few seeds with irregular embryo morphology were found. No abnormalities in endosperm development were found in *Candidate 14* (Col-0 background) seeds, and no difference was observed in the timing of endosperm cellularisation compared to Col-0 (Figure 5.9d,e).

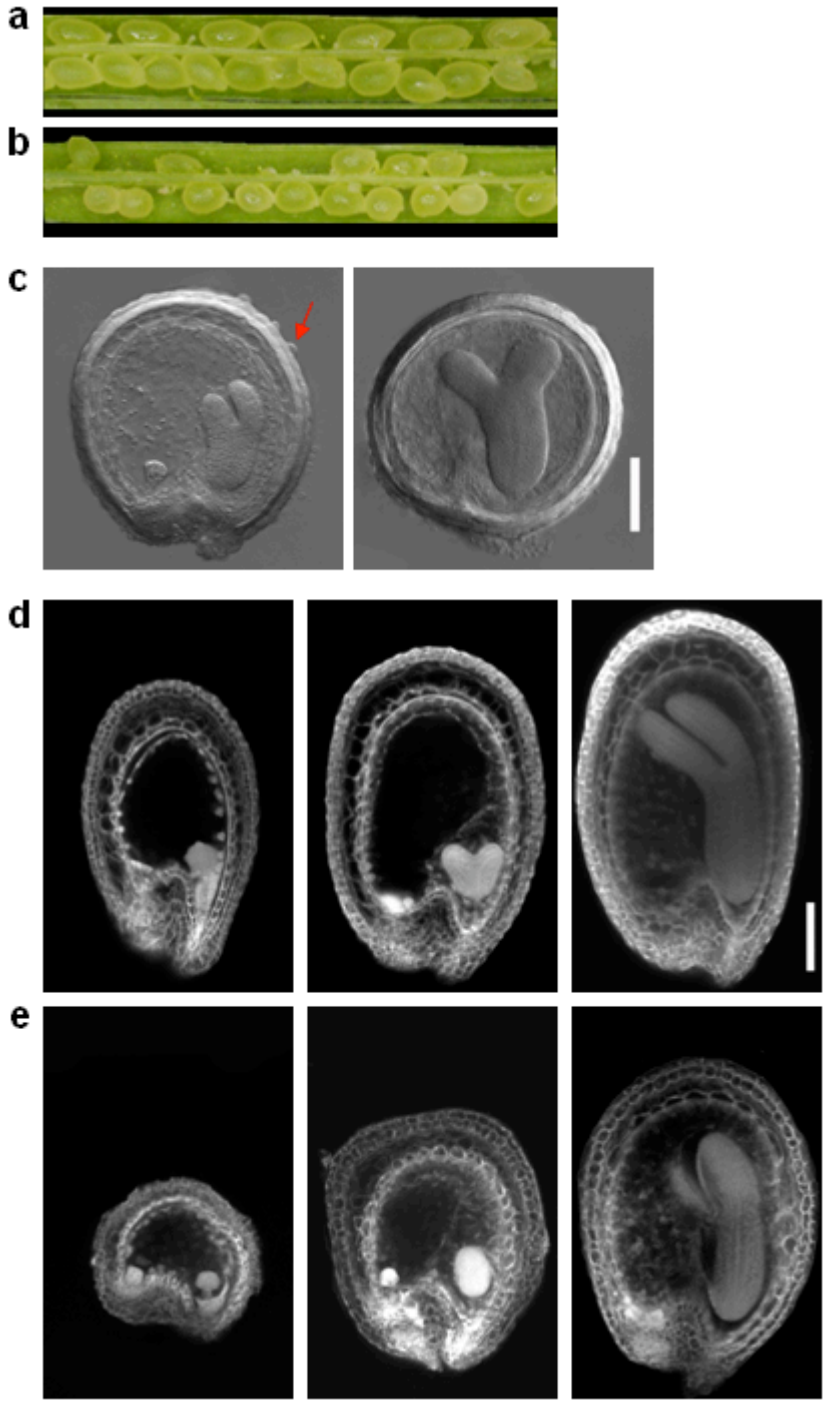

Figure 5.9: Seed phenotype of *Candidate 14* (Col-0 background).

Open siliques of Col-0 (a) and *Candidate 14* in the Col-0 background (b). (c) Cleared seeds imaged with differential contrast optics from *Candidate 14* in the Col-0 background. Red arrow = abnormal epidermal cells. Confocal images of Col-0 (d) and *Candidate 14* in the Col-0 background (e) at 3 DAP (left), 5 DAP (centre) and 7 DAP (right). Bar = 100 µm.

5.2.5 The Candidate 14 phenotype is caused by a mutation in the ARABIDOPSIS CRINKLY 4 (ACR4) gene.

High resolution gene mapping is a long and time consuming process. Therefore prior to commencing this work, a thorough search of the literature was carried out to determine whether the seed phenotype observed in *Candidate 14* had been previously reported (Table 5.1). This analysis suggested that the phenotype observed in *Candidate 14* bore a striking resemblance to that of *acr4*. The *Arabidopsis* mutant *acr4* produces seeds that are round and rough in appearance rather than elliptical and smooth as in the wild-type (Gifford *et al*., 2003; Watanabe *et al*., 2004). All *acr4* ovules and developing seeds showed epidermal irregularities, including cells of abnormal size and shape, and occasionally of inappropriate types such as stomata. Some ovules lacked a recognisable embryo sac and 40-85% of ovules were unfertilised or the seeds aborted. Although Gifford *et al*. (2003) found no embryo abnormalities in *acr4* mutants; malformed embryos were found in transgenic *Arabidopsis* plants containing antisense *ACR4* transgenes (Tanaka *et al*., 2002). Although no major defects in sepal morphology were reported in *acr4*, cells at the sepal boundaries were disorganised compared to wild-type (Gifford *et al*., 2003).

Table 5.1: Documented *Arabidopsis* mutants with round seeds.

Mutant	Similarities to *Candidate 14*	Differences from *Candidate 14*	References
arabidopsis crinkly 4 (*acr4*)	Round seeds, abnormal cells in epidermal seed coat layer, unfertilised ovules, abnormal embryo morphology.	-	Tanaka *et al*., 2002 Gifford *et al*., 2003 Watanabe *et al*., 2004
aberrant testa shape (*ats*)	Round (heart-shaped) seeds.	*ats* integuments consist of only 3 rather than 5 cell layers.	Léon-Kloosterziel *et al*., 1994
apetala 2 (*ap2*)	Seeds have irregular shape but are generally rounder than wild-type.	Striking floral phenotype.	Jofuku *et al*., 1994 Jofuku *et al*., 2005 Ohto *et al*., 2005
'ugly' in Chinese (*tso1*)	Round seeds (only weak alleles produce seeds).	Very low fertility (only 1-3 seeds per silique), abnormal flowers.	Hauser *et al*., 1998
haiku1 (*iku1*), *iku2*, *miniseed 3* (*mini3*)	Slightly rounder than wild-type seeds.	Early endosperm cellularisation.	Garcia *et al*., 2003 Luo *et al*., 2005

In order to determine whether the *Candidate 14* seed phenotype was caused by a mutation in the *ACR4* (At3g59420) gene, three tests were carried out:

1. Comparison of the seed phenotype of *Candidate 14*, in the *arf2-8* and Col-0 background, and those of the previously studied *acr4-2* mutant (Gifford *et al*., 2003) (Figure 5.10).
2. Allelism test between *Candidate 14* (Col-0 background) and *acr4-2*.
3. The *ACR4* gene was fully sequenced in *Candidate 14* (Figure 5.11).

The seed weight and seed shape of *Candidate 14* (Col-0 background) was compared to *acr4-2*. *Candidate 14* (Col-0 background) seeds were similar in weight to *acr4-2* seeds, and both were approximately 40% heavier than wild-type seeds (Figure 5.10b). In contrast, *Candidate 14* (*arf2-8* background) seeds were similar in weight to *arf2-8* mutant seeds. The seed length:width ratio was calculated to quantify seed shape, with a low ratio representing round seeds and a high ratio representing pointed seeds. The seed length:width ratio of *Candidate 14* seeds, whether in the *arf2-8* or Col-0 background, was not significantly different from *acr4-2* seeds (ANOVA followed by Tukey's multiple comparisons, $P = 0.518$), indicating that the seeds were of similar shape (Figure 5.10c). This is consistent with the hypothesis that the *Candidate 14* seed phenotype was the result of a mutation in the *ACR4* gene.

Candidate 14 (Col-0 background) was crossed with *acr4-2* to test for a disruption in the same gene. Six F_1 plants from each reciprocal cross, *Candidate 14* X *acr4-2* and *acr4-2* X *Candidate 14*, were visually screened by opening siliques. All F_1 progeny produced seeds that were round and rough in appearance. As the *Candidate 14* mutant phenotype was not lost in the F_1 plants, this allelism test provided additional evidence that *Candidate 14* contained a mutated version of the *ACR4* gene.

Genomic DNA was sequenced from *Candidate 14* mutants for the *ACR4* coding region plus 43 bases of the 5' and 114 bases of the 3' flanking regions (Figure 5.11). This region was sequenced as five fragments (primers and PCR conditions are detailed in Table 2.3 and Table 2.4). A single base change from G to A at position 33 from the translational start was identified in *Candidate 14* compared to the wild-type Col-0 sequence (www.arabidopsis.org). The *ACR4* coding region is not interrupted by introns; therefore the mutation had occurred within the coding region of the gene. The single base change introduces a premature stop codon, TGA (UGA in RNA), suggesting that a truncated protein of only 10 amino acids is produced in *Candidate 14* plants.

The above evidence confirms that the phenotype observed in *Candidate 14* is caused by a mutation in the *ACR4* gene. Therefore, from this point onwards *Candidate 14* (*arf2-8* background) will be referred to as *arf2-8 acr4* double mutants.

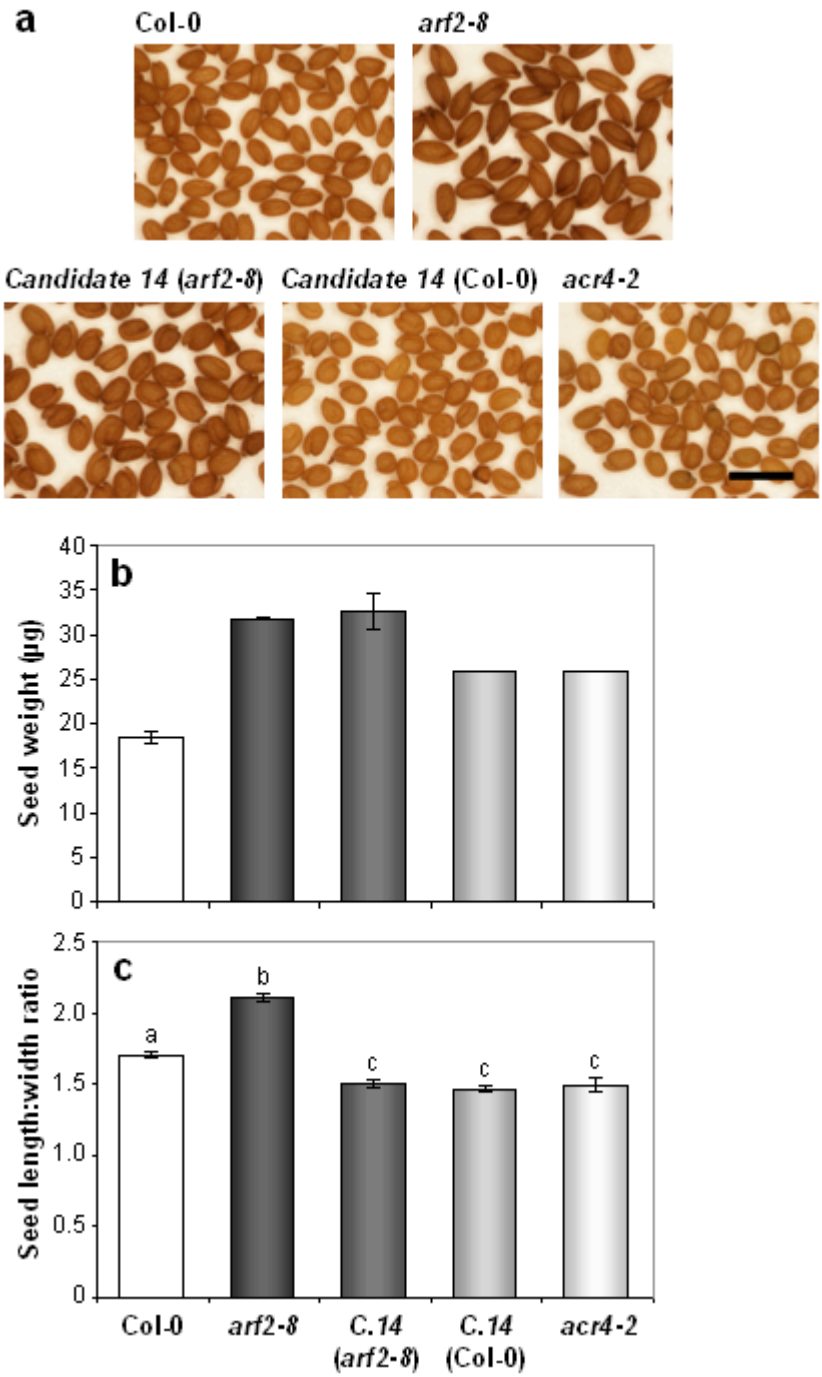

Figure 5.10: Comparing *Candidate 14* (Col-0 background) and *acr4-2* seeds.

(a) Photographs of Col-0, *arf2-8*, *Candidate 14* (*arf2-8* background), *Candidate 14* (Col background) and *acr4-2* seeds. Bar = 1 mm. (b) Comparison of seed weight. n ≥ 1 batch of 50 seeds. *acr4-2* seeds were kindly provided by G.C. Ingram, University of Edinburgh hence were not grown under the same conditions. (c) Comparison of the seed length:width ratio. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. n ≥ 10 seeds. Error bars = s.e.m.

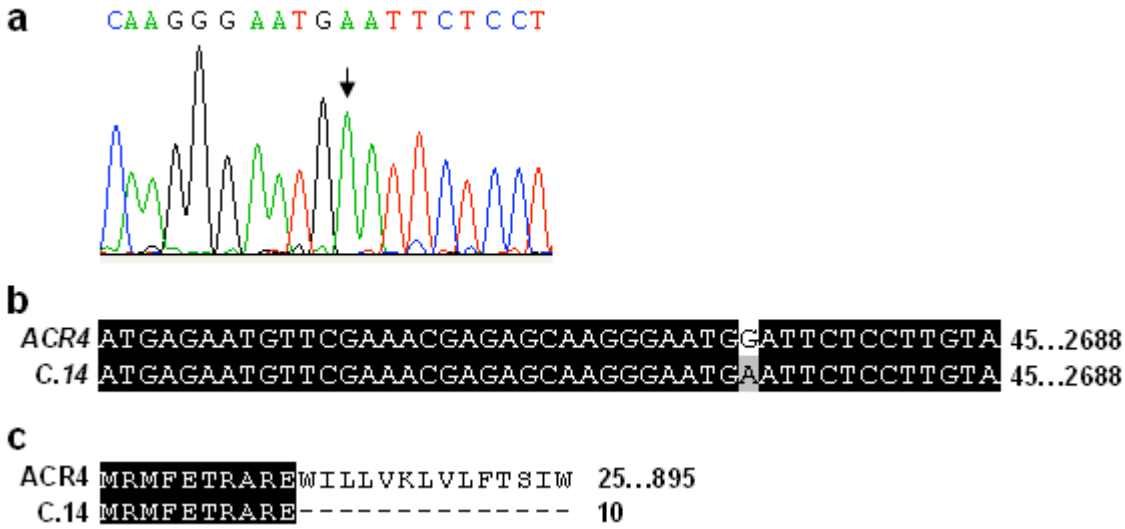

Figure 5.11: Sequencing of *ACR4* in *Candidate 14*.

(a) Chromatogram of *ACR4* sequence data obtained from *Candidate 14*. Arrow indicates the altered base. (b) Alignment of a 45 bp fragment of wild-type *ACR4* coding sequence with the mutated region in *Candidate 14* showing a single base change. The mutation is located 33 bp from the ATG and forms a premature stop codon. (c) Alignment of the N-terminal portion of the wild-type ACR4 and *Candidate 14* mutant proteins.

5.2.6 Seed shape in arf2-8 acr4 double mutants

With the aim of elucidating the mechanism by which *acr4* mutations alter *arf2-8* seed shape, the number of cells within four layers of the seed coat was counted in Col-0, *acr4-2* and *arf2-8* single mutants, and the *arf2-8 acr4* double mutant seeds at 5 DAP (Figure 5.12). Whilst *acr4* mutants have been previously shown to produce seed coat cells of altered shape and size (Gifford *et al*., 2003; Watanabe *et al*., 2004), the number of cells in each layer was not previously determined. This analysis showed that *acr4-2* seeds contained a similar number of cells in the seed coat to wild-type Col-0 seeds except in the oi1 layer, in which significantly fewer cells were present (Student's *t*-test, Col-0 vs. *acr4-2*, $P < 0.001$). As described earlier for *arf2-9* (Figure 4.5), developing seeds of *arf2-8* homozygous mutants contained significantly more cells in all seed coat layers measured than the wild-type post-fertilisation. In contrast, *arf2-8 acr4* double mutants contained significantly more cells than wild-type in only the ii1' layer of the seed coat, although this was significantly less than in the *arf2-8* mutant. Importantly, *arf2-8 acr4* double mutants had significantly more cells in three of the seed coat layers than *acr4-2* mutants alone. This suggests that *arf2-8 acr4* double mutants have an additive effect on integument/seed coat cell proliferation.

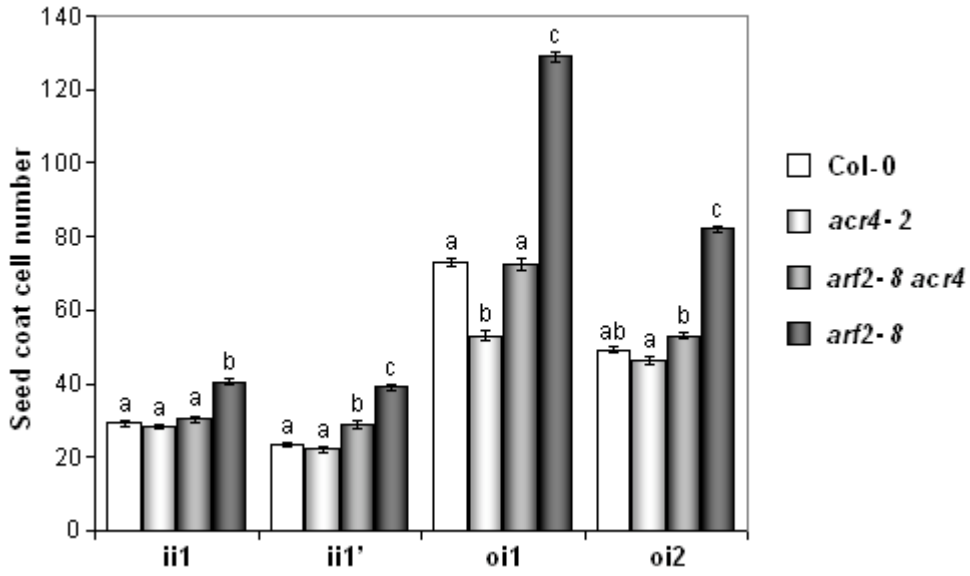

Figure 5.12: Seed coat development in *arf2-8 acr4* double mutants.

Comparison of the number of cells in the seed coat of Col-0, *acr4-2* and *arf2-8* single mutants, and the *arf2-8 acr4* double mutant seeds at 5 DAP. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. Error bars = s.e.m. n = 12.

In order to determine why seeds of *acr4-2* single mutants and *arf2-8 acr4* double mutants were round, the shape of individual cells of the seed coat was assessed in Col-0, *acr4-2*, *arf2-8* and *arf2-8 acr4* developing seeds. This analysis should also ascertain whether seed coat cell shape is associated with overall seed shape. For this analysis, cells of the ii1' and oi2 layers only were measured in seeds containing early heart stage embryos (Figure 5.13a). Three measurements were taken for each cell: length of the outside edge; length of the inside edge; and cell width (if the cell layer had expanded as often occurs in the ii1' layer of *arf2* seeds (Schruff *et al*., 2006), the total width of the layer was measured). In wild-type seeds, it was hypothesised that cells on the curved portion of the seed coat would have a greater difference between the length of the outside and inside edge of the cell compared to cells on a straight portion of the seed coat. Seeds were therefore partitioned into 'side' and 'curve' regions. The base of the seeds (below the tip of the adaxial ridge) was ignored; the lower two thirds of the remainder represented the 'side' while the upper third represented the 'curve'. As the ii1' layer only partially surrounds the embryo sac, measurements for the 'side' referred only to cells adjacent to the chalazal endosperm. In contrast, the difference between the outside and inside edge of the cells of rounder seeds, such as in *acr4* single and *arf2-8 acr4* double mutants, would be expected to be similar on the 'side' and on the 'curve'.

When examining the difference between the outside and inside cell lengths in the oi2 layer, there was no significant difference between wild-type and mutant seeds on both the 'side' and the 'curve' (Figure 5.13b). This indicated that there was no trend linking the shape of the cells in the oi2 layer and the shape of the seed. One possible explanation for this is that the width of the oi2 layer is small compared to the ii1' layer (Figure 5.13c). It is likely that a thinner layer will show a smaller difference between the outside and inside cell lengths, even on a curved edge.

In contrast to the oi2 layer, the width of the ii1' layer was greater and the cells showed more genotypic variation in shape, suggesting that this layer is associated with alterations in overall seed shape (Figure 5.13b,c). It was previously hypothesised that in longer seeds, such as in wild-type and *arf2-8* mutants, cells on the 'curve' would have a greater difference between the outside and inside cell length compared to cells on the 'side'. This was confirmed in both wild-type and *arf2-8* seeds as the ii1' cells on the 'curve' had a significantly larger difference between the outside and inside cell lengths from ii1' cells on the 'side' (Student's *t*-test, wild-type, P = 0.007; *arf2-8*, P = 0.012). As expected, in the round seeds of the *arf2-8 acr4* double mutant, the ii1' cells on the 'curve' did not show a significant difference between the outside and inside cell lengths from ii1' cells on the 'side' (Student's *t*-test, *arf2-8 acr4*, P = 0.096). This indicated that in these round seeds ii1' cells on the 'curve' were similar in shape to ii1' cells on the 'side'. In the round seeds of *acr4-2* mutants, cell shape did not seem to fit the expected trend. ii1' cells on the 'curve' had a significantly larger difference between the outside and inside cell lengths than ii1' cells on the 'side' (Student's *t*-test, *acr4-2*, P = 0.009), suggesting that *acr4-2* seeds have distinctive side and curve regions.

To simplify the interpretation of these results, the ratio of the difference between the outside and inside cell lengths of ii1' cells on the 'side' to the 'curve' was calculated. Thus, the side:curve ratio should be higher in long seeds than in round seeds, which should have a ratio of approximately one since if entirely round there should be no distinction between the 'side' and the 'curve'. In accordance with this, wild-type seeds had a ratio of 2.08, the even longer seeds of *arf2-8* a ratio of 2.84, and the round seeds of the *arf2-8 acr4* double mutant had a ratio of 1.33. As mentioned above, ii1' cells of *acr4-2* seeds did not behave in the predicted manner, and had a ratio of 2.65 which was similar to the ratio of the long *arf2-8* seeds. The low side:curve ratio of ii1' cells in *arf2-8 acr4* double mutant seeds had occurred as a result of a reduced difference between the outside and inside cell lengths on the 'curve' rather than an increase on the 'side'. This is surprising since the rounder shape of *arf2-8 acr4* seeds would suggest that the 'side' had become more curved, not that the 'curve' had become straighter.

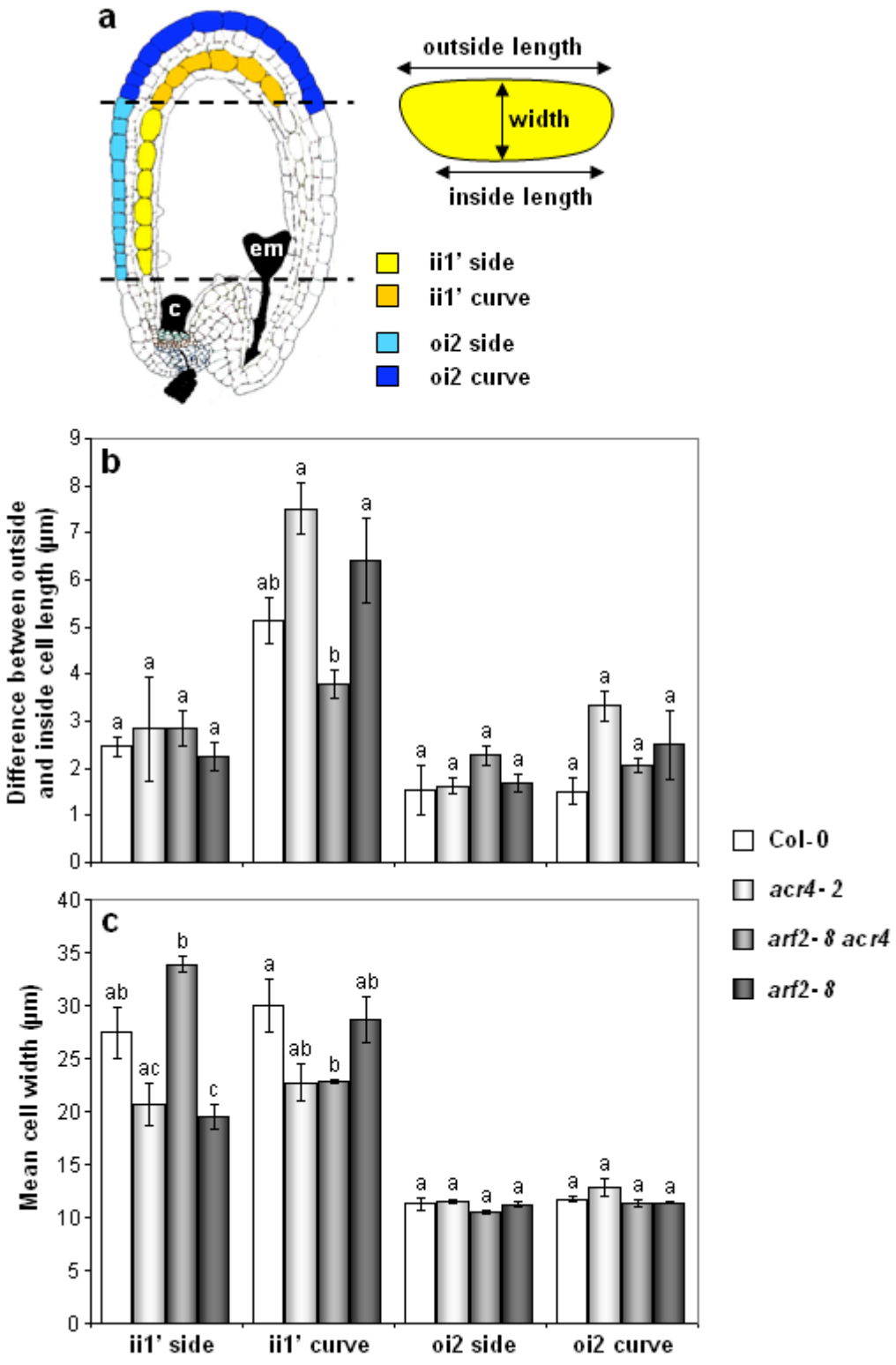

Figure 5.13: Does seed coat cell shape determine overall seed shape?

(a) Diagram to demonstrate how seed coat cell dimensions were measured. em = embryo. c = chalazal endosperm. Adapted from Debeaujon *et al*., (2003) (b) Comparison of the difference between the outside and inside cell length of two layers of the integument in heart stage embryos of Col-0, *acr4-2*, the *arf2-8 acr4* double mutant and *arf2-8*. (c) Comparison of cell width in two layers of the integument in heart stage embryos of Col-0, *acr4-2*, the *arf2-8 acr4* double mutant and *arf2-8*. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. Error bars = s.e.m. n ≥ 3 seeds.

5.2.7 Candidates 19 and 20: suppressors of the arf2 seed phenotype

Candidate 19 and *Candidate 20* have been presented together as they displayed similar phenotypes. It is possible that these two candidates had been caused by an EMS-induced mutation in the same gene yet this is unlikely as they were obtained from separate M_2 families. The mature seeds of *Candidates 19* and *20* were rounder than those of *arf2-8* hence both lines were considered *arf2* suppressors (Figure 5.3). Developing seeds in young siliques of *Candidates 19* and *20* were also easily distinguishable from *arf2-8* seeds due to their round shape (Figure 5.14a,b,c) The round seed phenotype was present in the M_3 and M_4 plants indicating that the candidate phenotypes were the result of EMS-induced mutations which caused a heritable change in seed shape. Although *Candidates 19* and *20* had altered seed shape, seed weight was not significantly different from *arf2-8* in either of the lines (Student's *t*-test, *Candidate 19* vs. *arf2-8*, P = 0.910; *Candidate 20* vs. *arf2-8*, P = 0.468) (Figure 5.14d). Seed clearing revealed no obvious alterations to seed coat development in either *Candidates 19* or *20* that could be responsible for the change in seed shape (Figure 5.14e,f,g,h). The adaxial ridge is enlarged in both *Candidates 19* and *20*, as in *arf2-8* mutant seeds. *Candidates 19* and *20* were similar to that of *arf2-8* in rosette (Figure 5.4) and floral morphology.

Candidates 19 and *20* were backcrossed to *arf2-8* to determine the segregation ratio of the seed phenotypes and ascertain whether they were the result of a recessive or dominant mutation. The F_2 population resulting from the *Candidate 19* X *arf2-8* contained fifteen plants with round seeds and thirty-one plants with *arf2-8*-like seeds (For a 3:1 segregation, $\chi^2 = 1.420$, P = 0.233). The F_2 population resulting from the *Candidate 20* X *arf2-8* contained twenty-one plants with round seeds and fifty-seven plants with *arf2-8*-like seeds (For a 3:1 segregation, $\chi^2 = 0.154$, P = 0.695). Therefore, the seed phenotype of *Candidates 19* and *20* were associated with single gene recessive mutations.

To determine whether *Candidates 19* and *20* displayed visible seed phenotypes in a wild-type background, the candidate mutants were also crossed to Col-0. In the F_2 population, it was difficult to distinguish the seed phenotypes of the candidates from wild-type seeds by visual comparison alone (Figure 5.15a). The length and width of the seeds from individual plants of the *Candidate 19* X Col-0 F_2 population was measured to determine whether some plants produced seeds with altered shape. The length:width ratios of several plants were found to be significantly lower than in Col-0, suggesting that *Candidate 19* did have a phenotype in a wild-type background (Figure 5.15b).

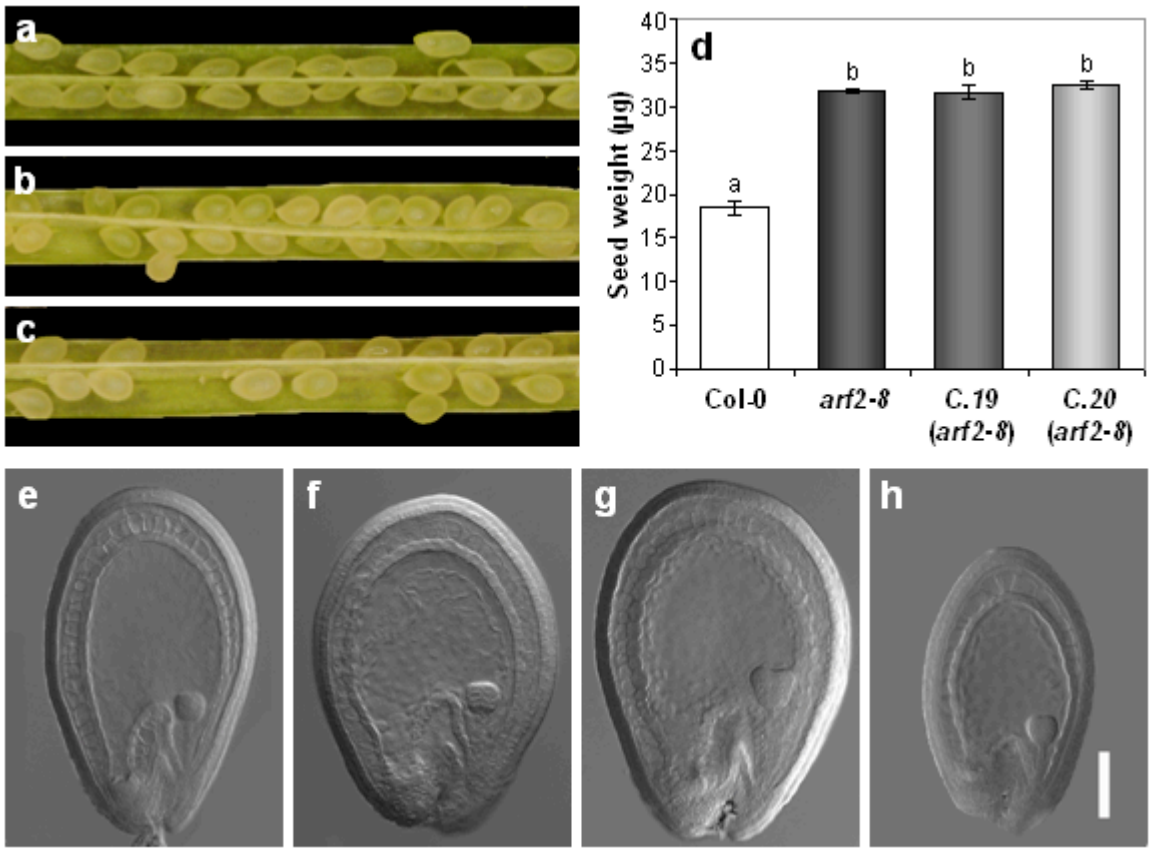

Figure 5.14: Seed phenotype of *Candidates 19* and *20* (*arf2-8* background).

Open siliques of *arf2-8* (a), *Candidate 19* in the *arf2-8* background (b) and *Candidate 20* in the *arf2-8* background (c). (d) Comparison of mean seed weight. Error bars = s.e.m. n ≥ 3 batches of 50 seeds. Seed weight values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. Cleared seeds imaged with differential contrast optics from *arf2-8* (e), *Candidate 19* in the *arf2-8* background (f), *Candidate 20* in the *arf2-8* background (g) and Col-0 (h). Photographs were taken on different occasions and are not of seeds of the same age; however they are of the same magnification and are given to demonstrate seed shape and seed coat morphology only. Bar = 100 µm.

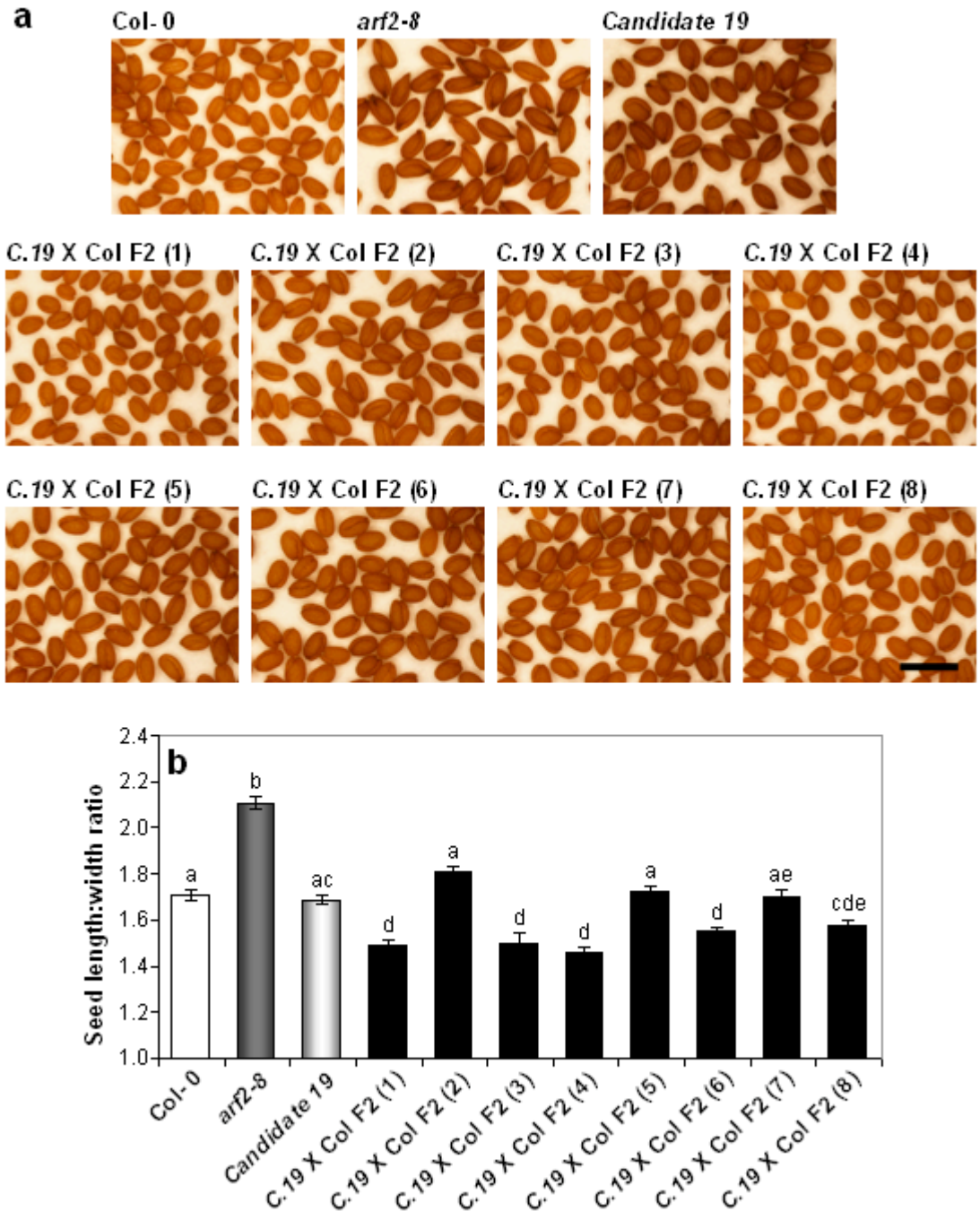

Figure 5.15: Seed phenotype of *Candidates 19* (Col-0 background).

(a) Photographs of mature seeds from Col-0, *arf2-8*, *Candidate 19* (*arf2-8* background), and eight plants from the *Candidate 19* X Col-0 F_2 population. Bar = 1 mm. (b) Comparison of seed length:width ratio. Error bars = s.e.m. $n \geq 11$. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters.

In order to map the gene responsible for the seed phenotype of either *Candidate 19* or *20*, an F_2 mapping population must first be generated. Within this population, it must be possible to easily distinguish the candidate mutant phenotype. The seed phenotypes of *Candidates 19* and *20* were more clearly visible in an *arf2-8* mutant background than in a wild-type background; hence all plants in the F_2 mapping population should be homozygous for *arf2-8*. In addition, the *arf2-8* pointed seed phenotype was more prominent in the L*ER* rather than L*er* background (see section 5.2.8). Therefore, *Candidates 19* and *20* were crossed to L*ER* (*arf2-8* background) in order to generate an F_2 mapping population. The ratio of round seeds to pointed seeds in the F_2 mapping population was subsequently determined to ensure that mutants could be distinguished (For a 3:1 ratio: *Candidate 19*, 16 round, 31 pointed, $\chi^2 = 2.050$, P = 0.152; *Candidate 20*, 14 round, 29 pointed, $\chi^2 = 1.310$, P = 0.252).

5.2.8 arf2-8 in a Landsberg background

High resolution positional (map-based) cloning requires a high density of genetic markers. Col-0 and L*er* are sufficiently divergent, differing at 4-11 positions in every 1 kb, to support the design of molecular markers and as extensive sequence information and pre-designed markers are available, are the most commonly used ecotypes for high resolution mapping (Lukowitz *et al*., 2000). If the phenotype of a particular candidate cannot be visualised in a wild-type background, it will be necessary to carry out gene mapping in the *arf2-8* mutant background. The candidate lines were generated in a Col-0 background containing the *arf2-8* mutation. It is therefore necessary to cross these plants with *arf2* mutants in the L*er* background. However, despite reports of nine different *arf2* mutant alleles to date (Li *et al*., 2004; Okushima *et al*., 2005; Schruff *et al*., 2006); no *arf2* mutants have been identified in the L*er* background. Consequently, *arf2-8* mutants (L*er* background) were generated by backcrossing *arf2-8* (Col-0 background) into L*er* eight times. In addition, *arf2-8* (L*er*) plants were crossed with L*ER* to remove the *erecta* mutation.

The effect of *ARF2* loss-of-function mutations on plant growth and development have not previously been described in the L*er* or L*ER* backgrounds. Therefore, *arf2-8* (L*er*) and *arf2-8* (L*ER*) plants were analysed for several aspects, particularly those related to seed size and yield (Table 5.2). *arf2-8* mutant plants whether in the L*er* or L*ER* background had a dramatically lower seed weight than *arf2-8* mutants in the Col-0 background. This reduction in seed size appears largely due to improved fertility, as *arf2-8* (L*er*/L*ER*) set more seeds per silique and had a higher percentage of elongated siliques than *arf2-8* (Col-0). It was also observed that *arf2-8* (L*er*/L*ER*) seeds were not as pointed

as *arf2-8* (Col-0) seeds (Figure 5.16a), and this was confirmed by the reduced length:width ratio of these seeds (Figure 5.16b). However, the length:width ratio of *arf2-8* mutants in the L*ER* background was considerably higher than in the L*er* background. This indicates that the L*ER* background would be a better choice for the mapping cross, as round seed phenotypes of the candidate mutants would be more easily distinguished from the *arf2-8* seed phenotype.

Table 5.2: Comparison of *arf2-8* in Col-0, L*er* and L*ER* genetic backgrounds.

Genotype	Mean seed weight (µg)	Number of seeds per silique	Percentage of elongated siliques
Col-0	19.63 ± 0.44^{a}	58.3 ± 2.7	96.7 ± 0.6
L*er*	17.37 ± 0.11^{b}	67.3 ± 3.7	100.0 ± 0.0
L*ER*	18.97 ± 0.29^{ab}	70.5 ± 0.6	99.4 ± 0.6
arf2-8 (Col-0)	32.68 ± 0.10^{c}	10.7 ± 10.2	10.0 ± 1.9
arf2-8 (L*er*)	23.10 ± 0.59^{d}	29.4 ± 7.9	75.8 ± 3.2
arf2-8 (L*ER*)	24.09 ± 0.55^{d}	ND	77.6 ± 2.3

Mean ± s.e.m. $n \geq 3$ plants. Seed weight values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different superscript letters. Number of seeds per silique was determined on the 20^{th} silique on the primary inflorescence. The percentage of elongated siliques referred to those on the primary silique only. ND = not determined.

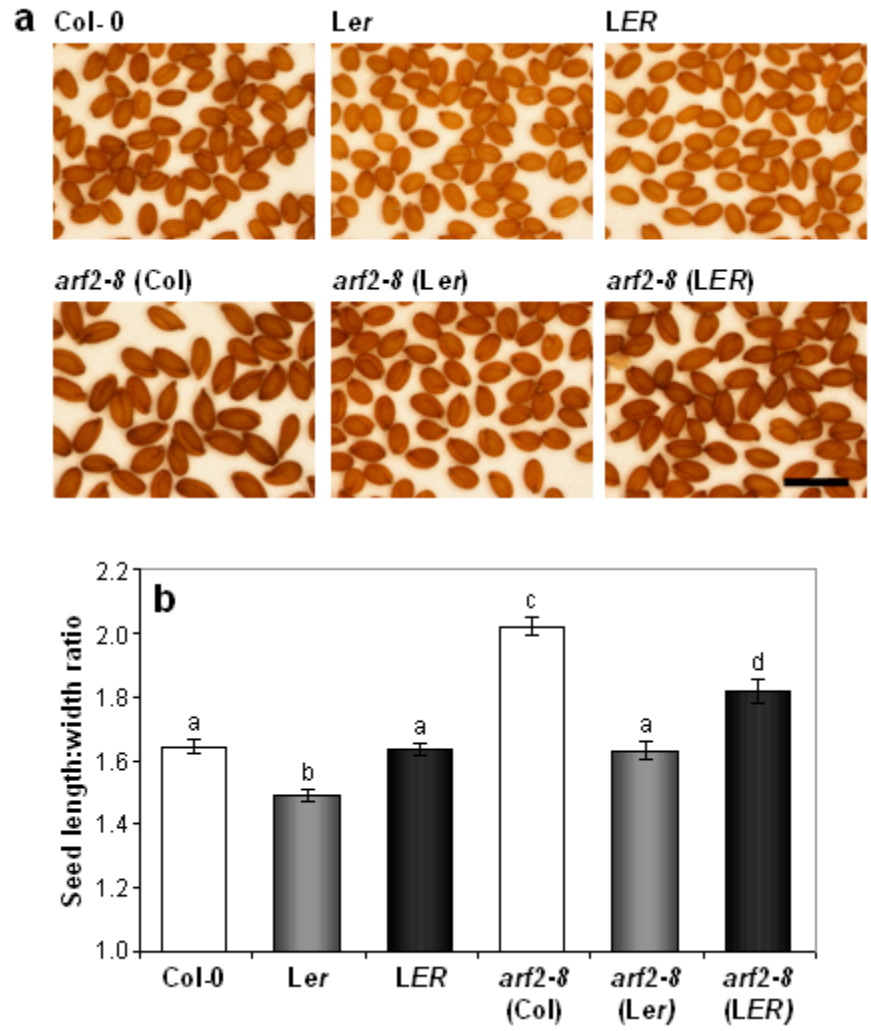

Figure 5.16: Seed phenotype of *arf2-8* in Col, L*er* and L*ER* genetic backgrounds.

(a) Photographs of Col-0, L*er*, L*ER*, *arf2-8* (Col), *arf2-8* (L*er*) and *arf2-8* (L*ER*) seeds. Bar = 1 mm. (b) Comparison of the seed length:width ratio. n = 20. Error bars = s.e.m. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters.

5.3 Discussion

The primary aim of the work reported in this chapter was to elucidate novel genes involved in the regulation of seed size and shape in *Arabidopsis*. EMS mutagenesis was employed to induce second-site mutations in *arf2-8* mutants, and thereby uncover suppressors and enhancers which have the potential to reveal more information about the role of *arf2* in seed development.

5.3.1 How is seed shape determined in Arabidopsis?

In a screen for suppressors and enhancers of the *arf2-8* seed phenotype, the *acr4* gene was identified as an *arf2* suppressor. *arf2-8 acr4* double mutants produced round seeds rather than the distinctive pointed seeds of *arf2-8* mutants alone. Sequencing of the *ACR4* gene revealed that a single base pair mutation had been induced 33 bp from the translational start, which led to the production of a premature stop codon that is likely to cause a complete loss of *ACR4* function. To determine the mechanism by which loss of *ACR4* function altered *arf2-8* seed shape, the number of cells in the seed coat was determined (Figure 5.12). In *arf2-8 acr4* double mutants, an additive effect on cell proliferation within the seed coat was observed, as they contained an intermediate number of cells compared to *acr4-2* and *arf2-8* mutants alone. This indicated that the *acr4* mutation affects the seed coat through a developmental pathway that is independent of *arf2*. Previous research examining the *aberrant testa shape* (*ats*) seed shape mutant also demonstrated an additive effect when combined with mutations in *glabrous 2* (*gl2*), *ttg* and *ap2* (Léon-Kloosterziel *et al*., 1994). The present research therefore supports the proposal made by Léon-Kloosterziel *et al*. (1994) that seed shape and thus integument/seed coat development is a complex process involving many genes that control cell division and cell differentiation.

Despite the fact that in mature *Arabidopsis* seeds the embryo fills the seed completely, it is the seed coat rather than the embryo that determines seed shape (Léon-Kloosterziel *et al*., 1994). Similarly, it is the integument/seed coat and not the endosperm that controls seed shape in young developing seeds. This is apparent since mutants with altered seed shape such as *arf2*, *acr4*, *ats*, and *ap2* have abnormal integument/seed coat development (Schruff *et al*., 2006; Gifford *et al*., 2003; Léon-Kloosterziel *et al*., 1994; Jofuku *et al*., 1994). In contrast, mutants that effect endosperm development such as *iku1*, *iku2* and *mini3* tend to show modifications in seed size with little or no change to seed shape (Garcia *et al*., 2003; Luo *et al*., 2005). Additional evidence for the important role of the integuments/seed coat in determining seed shape is available as this component of the seed is entirely derived from maternal tissues. Mutants in testa development have a

maternal effect on seed shape, thus when *arf2-8* mutants are crossed with a wild-type pollen parent, large pointed seeds are produced even though the *arf2* mutation is recessive (Schruff *et al*., 2006).

The integuments/seed coat clearly have a crucial role in determining seed shape, but how they achieve this is at present not well understood. We measured the shape of individual cells in the seed coat to determine whether cell shape affected overall seed shape (Figure 5.13). One important finding was that the shape of cells in the ii1' layer was linked to seed shape to a greater extent than the shape of cells in the oi2 layer. It is likely that this was the case as cells of the ii1' layer had a greater width. A second key finding was that while cells in the ii1' layer of wild-type, *arf2-8* single mutants and *arf2-8 acr4* double mutants were the shape predicted by our model for both long and round seeds, those of *acr4-2* single mutants were not. It was hypothesised that the curve:side ratio (Section 5.2.6) would be close to one for round seeds and considerably higher for long seeds. The curve:side ratio for the round seeds of *acr4-2* mutants was 2.65, suggesting that these seeds were longer than wild-type seeds, which had a curve:side ratio of 2.08.

Functional ACR4 has been shown to be required for correct cellular organisation during ovule integument outgrowth since *acr4* mutant integuments and seed coats contained cells of abnormal shapes and sizes, but also occasionally of inappropriate cell types such as stomata (Gifford *et al*., 2003). It is therefore possible that *acr4-2* mutant seeds did not fit our model as cells of the ii1' layer were irregularly shaped. Gifford *et al*. (2003) observed that *acr4* mutant integuments were thicker than those of wild-type. This suggests that the integument cells were undergoing cell expansion, but that this expansion was not restricted to one direction. While in wild-type seeds directional cell elongation results in lengthening of the abaxial integuments, in *acr4* the integuments failed to elongate sufficiently to give the curvature of wild-type seeds. Since most seed coat cell layers contained the same number of cells in *acr4* as in wild-type (Figure 5.12), the shorter seed coat must have resulted from a lack of directional cell elongation rather than a decrease in the number of integument cells. In *arf2-8 acr4* double mutants, a significant increase in seed coat cell number compared to *acr4* mutants alone enabled the formation of a longer seed coat, but due to a lack of cell identity caused by loss-of-function of *ACR4* the seeds remained round.

5.3.2 Further Work

This screen successfully led to the identification of several candidate mutants classified as either *arf2-8* suppressors or enhancers according to their seed phenotype. We have shown that *Candidate 14* was caused by a mutation within *ACR4*. The analyses described show

that it would also be possible to identify the mutated genes responsible for other candidate lines. *Candidate 4*, an *arf2-8* enhancer, produced seeds that resembled those of *ttg2* in a Col-0 background. An allelism test and sequencing should be used to assess whether *Candidate 4* was caused by a mutation in the *ttg2* gene. If this is not the case, gene mapping could be used to identify the mutated gene since the *Candidate 4* seed phenotype was associated with a single gene recessive mutation and was clearly visible in a Col-0 background. The genes responsible for the seed phenotypes of *Candidates 19* and *20* could also be identified though gene mapping; however this would need to be carried out in an *arf2-8* background since the seed phenotype was difficult to distinguish in a Col-0 background.

Extensive backcrossing of *arf2-8* into the L*ER*/L*er* background revealed that *arf2-8* mutants were more fertile and produced rounder seeds in the L*ER*/L*er* background compared to the Col-0 background. Increased fertility of *arf2*-8 (L*ER*/L*er* background) was associated with a higher seed set per silique and a greater number of elongated siliques. To determine how fertility was improved the floral organ lengths of *arf2-8* (L*ER*/L*er* background) should be determined since over-elongation of the sepals and gynoecia were shown to reduce self-pollination in the Col-0 background (Schruff *et al*., 2006). To establish why *arf2-8* seeds were not as pointed in the L*ER*/L*er* as in the Col-0 background, the number of cells in the seed coat should be determined.

Using *arf2-8* and *acr4-2* single mutants along with the *arf2-8 acr4* double mutant, an initial investigation to determine whether the shape of seed coat cells was associated with overall seed shape was carried out. Further analysis is necessary to fully understand this relationship, especially since *acr4-2* seeds did not fit the predicted model. This analysis could include establishing: the thickness of the entire seed coat; the mean area of cells within each seed coat layer; and the mean cell length. An extension of this work could include studying the seed coats of other mutants with altered seed shape.

Through this screen, *ACR4* was identified as an *arf2-8* suppressor. *arf2-8 acr4* double mutants exhibit an additive effect on seed development suggesting that these two genes are both involved with development of the integument/seed coat but act through independent developmental pathways. As an alternative method to identify proteins that interact directly with *arf2*, a yeast 2-hybrid screen could be conducted.

5.3.3 Summary

Four suppressors and one enhancer of the *arf2* seed phenotype were investigated in detail:

Candidate 3 (suppressor) produced rounder seeds than *arf2-8*, a phenotype associated with abnormal development of the ii1' layer of the seed coat. In addition, *Candidate 3* mutants had only a few small rosette leaves, abnormal carpel development and reduced pollen production. Despite the interesting phenotype of *Candidate 3*, further work on this line was discontinued due to difficulties with crossing and complicated genetics.

Candidate 4 (enhancer) seeds were narrower and pointier than those of the *arf2-8* mutant alone. The *Candidate 4* seed phenotype was visible in a Col-0 background, with seeds appearing small and pale, and bearing resemblance to those of *transparent testa glabra 2* (*ttg2*) seeds (Johnson *et al*., 2002).

Candidate 14 (suppressor) produced round seeds, a phenotype found to be caused by a recessive mutation in the *ACR4* gene. In *arf2-8 acr4* double mutants, an additive effect on cell proliferation within the seed coat was observed, as they contained an intermediate number of cells compared to *acr4-2* and *arf2-8* mutants alone. This indicated that the *acr4* mutation affects the seed coat through a developmental pathway that is independent of *ARF2*.

Both *Candidates 19* and *20* (suppressors) produced rounder seeds than *arf2-8* mutants alone. The seed phenotypes of these candidates were caused by single gene recessive mutations. Evidence for a seed phenotype in a Col-0 background was observed for *Candidate 19*; however identification of the genes responsible for both *Candidates 19* and *20* would be easier in an *arf2-8* background.

6. USING *ARABIDOPSIS* AS A MODEL FOR YIELD AND YIELD COMPONENT ANALYSIS

6.1 Introduction

6.1.1 Yield component analysis in crop plants

Seed yield is an extremely complex trait. Hence it has frequently been divided into various components in order to simplify its study. Yield components have been utilised since the 1920s in a number of crop plant species (Egli, 1998). Breeders have often selected for increases in a particular yield component, such as seed number per plant or mean seed weight, with the aim of improving seed yield. Therefore, it is important to identify which components contribute most to yield. Given that yield in crops is measured as the seed weight per unit area, the two primary yield components are seed size and seed number. The number of seeds per unit area is greatly influenced by the environment, and thus is usually closely associated with seed yield. A positive relationship between seed number and yield has been shown in soybean, wheat (Egli, 1998), maize (Jong *et al*., 1982) and sunflower (Cantagallo *et al*., 1997). In comparison, mean seed weight is often not associated with yield (Peltonen-Sainio *et al*., 2007). Hence, a plant producing large seeds will not necessarily be high-yielding. There are numerous other factors which affect yield, including leaf number, leaf area, plant height, total dry mass of aerial organs, plant biomass, ovule number per fruit, fruit or flower number and flowering time. These factors can influence yield by altering either seed size, seed number, or both.

The popularity of studying yield components has fluctuated as breeders regularly encountered 'yield component compensation'. For instance, successful selection for increased seed size was accompanied by a reduction in seed number so that the seed yield remained unchanged (Hartwig and Edwards, 1970; White and Izquierdo, 1991). This trade-off between seed number and seed size is believed to exist as the yield components share a common limited supply of resources (Harper *et al*., 1970; Venable, 1992). Many examples of the negative correlation between seed number and seed size have been documented in crop species (Kiniry, 1988; Kiniry *et al*., 1990; Borrás and Otegui, 2001).

Yield is the end product of many physiological processes that are sensitive to the environment and occur throughout plant development. When the yield production process is considered over time it becomes even more complex; however as yield components can each be associated with a specific developmental stage some clarity can be achieved. One study in the field bean, *Phaseolus vulgaris*, considered three yield components and the

sequential order of their development: number of pods per plant, number of seeds per pod and the mean seed weight (Adams, 1967), and consequently suggested that seed number is determined before seed size. The sequential concept was further advanced by dividing the yield production process into three main stages (Murata, 1969):

1. Formation of organs for nutrient uptake and photosynthesis.
2. Formation of reproductive organs and the yield container.
3. Production, accumulation and translocation of yield contents.

While seed number and the maximum potential seed size are determined in the second stage, seed filling and thus how much of the yield potential is realised is determined in the last stage. It has been suggested that seed number rather than seed weight is more closely associated with yield, simply because it is determined first during crop development (Egli, 1998). However this concept is complicated as the stages are not always distinct, especially in plants with an indeterminate growth pattern, such as soybean, oilseed rape and *Arabidopsis*.

Although seed size is often not associated with yield, this yield component has received more attention from scientists than any other, probably because it is readily observable. Final seed size is determined by the extent of its two components: seed growth rate (SGR) and seed fill duration (SFD). Like the components of yield, both SGR and SFD are affected by environmental and genetic factors (Sadras and Egli, 2008). Variation in seed size is closely linked to SGR. For example, the small seeds of rice grow at a rate of 1.3 mg seed^{-1} day^{-1} whereas the larger seeds of maize grow faster at 7.4 mg seed^{-1} day^{-1} (Egli, 2006). However, SGR can also be influenced by the supply of resources to seeds which can be altered by shading or defoliation during the seed filling period. SFD is sensitive to temperature and water stress, yet does not respond consistently to nutrient supply (Egli, 2006).

Although seed size is considered a remarkably stable trait (Harper *et al*., 1970), dramatic changes in seed weight have been observed by altering the post-flowering source-sink ratio (Borrás and Otegui, 2001; Gambin and Borrás, 2007). Pollination treatments were used to artificially modify kernel number per plant in maize and obtain a range of source-sink ratios (Borrás and Otegui, 2001). In this experiment, a low kernel number (representing a high source-sink ratio) resulted in a high kernel weight. Kernel weight was found to be closely related to SGR but not to changes in SFD. Similar experiments have been carried out in other crop plants with varying results (Borrás *et al*., 2004). For instance, wheat grain size did not change considerably with an increase or decrease in assimilate availability, suggesting that this crop is primarily sink limited. Thus, the degree of plasticity in seed size varies in different crop plants.

6.1.2 Yield and the HI

The ratio of seed yield to total plant biomass, termed the HI, is an important trait associated with substantial increases in crop yields that occurred during the Green Revolution (Sinclair, 1998). Often yield improvements were obtained without any change to total plant biomass (Austin *et al*., 1980; Perry and D'Antuono, 1989), indicating that a greater proportion of assimilates had been partitioned to reproductive development. Therefore, with the same level of inputs such as fertilisers, water and agrochemicals, a greater seed yield was achieved in crop plants with a superior HI ratio.

During the Green Revolution, HI was significantly raised by the introduction of dwarf wheat and rice varieties. The dwarf cereal varieties were short due to abnormal responses to the growth-promoting hormone gibberellin (GA). The Green Revolution genes have recently been identified: the *REDUCED HEIGHT1* (*RHT-B1*/*RHT-D1*) gene in wheat and the *SLENDER RICE1* (*SLR1*) gene in rice (Peng *et al*., 1999; Ikeda *et al*., 2001). These genes are orthologous to maize *DWARF8* (*D8*), barley *SLENDER1* (*SLN1*) and *Arabidopsis GA INSENSITIVE* (*GAI*). Mutations in these genes are of two types; the semi-dominant gain-of-function mutations, *gai*, *d8* and *rht1*, lead to GA-insensitive dwarf phenotypes whereas the recessive loss-of-function mutations, *sln1* and *slr1* cause constitutive GA responses and elongated phenotypes. Understanding the genetic and physiological basis of high-yielding varieties is important for improving crop production. The modification of plant architecture had a positive impact on yields in the 20th-century and through biotechnology may continue to play an important role (Sakamoto and Matsuoka, 2004).

6.1.3 Yield on a larger scale

Yield may be studied at various levels: molecule, tissue, organ, plant and crop (Thornley, 1980). This thesis describes experiments which study yield at the molecular level (e.g. the affect of *ARF2* on yield), at the level of the tissue (e.g. integuments or endosperm), at the level of the organ (e.g. the seed), and at the level of the plant (e.g. seed number per plant, total seed yield and HI). However, economic yield is always measured on a land area basis and therefore is best studied within a plant community where important factors can be considered such as plant population, leaf area index and light interception (Egli, 1998). Within a population, plants compete for resources such as light, water and nutrients. At high planting densities, the competition for resources becomes more intense.

Unlike small grain crops, the HI of maize was already high at the beginning of the 20th century (Hay, 1995). However, substantial increases in maize yields have been achieved due to selection of hybrids that are adapted to high population densities

(Tollenaar and Wu, 1999). Due to the importance of planting density in crop fields, many studies have monitored the effect of different densities on seed yield, total biomass and HI. A model of the relationship between density, yield, biomass and HI has been proposed (Donald and Hamblin, 1976). Total biomass increases with density to a maximum value which is determined by the environmental conditions and at higher densities remains constant, provided lodging or other interfering factors are not present. Seed yield increases with density to a maximum value but declines steadily as density is increased further. The optimum seed yield and total biomass are often synchronised to the same density. Contrastingly, increasing density is associated with a decrease in the HI ratio.

6.1.4 Yield in Arabidopsis

A large research effort has been applied to seed yield and its components in crop plants, yet the molecular machinery controlling this complex trait is not fully understood. Although *Arabidopsis* has no direct economic importance, it can be used to elucidate the genes and regulatory networks controlling plant processes. *Arabidopsis* is currently being used to gain further insight into other agronomically-relevant traits, such as seed oil composition (O'Neill *et al*., 2003), heterosis (Barth *et al*., 2003) and biomass production (Gonzalez *et al*., 2008). The objective of the work reported in this chapter was to utilise the natural variation present in *Arabidopsis* to assess the variation in yield and the relationships between yield components in this species since this has not been previously determined.

More than 750 *Arabidopsis* natural populations, referred to as ecotypes or accessions, have been sampled from diverse environmental conditions from across the globe. The geographical origins of the ecotypes used in the present study are shown in Figure 6.1. Despite belonging to the same species, *Arabidopsis* ecotypes vary greatly in their morphology and development. This study aimed to establish the seed size, yield and HI of several *Arabidopsis* ecotypes and use this information to determine the relationship between yield components in this model species. Two genotypes that differ in plant stature have also been grown under different planting densities in order to determine how seed size, yield and HI are affected within a plant community. There is some evidence in maize to suggest that a short, compact stature can be advantageous at high densities (Nelson and Ohlrogge, 1957; Begna *et al*., 1997). The *Arabidopsis* ecotype L*er* is commonly used as a 'wild-type' but is actually an X-ray-induced mutant. The *erecta* (*er*) mutant has a compact inflorescence due to short internodes, clustered flower buds, short pedicels, round leaves, and short, blunt siliques (Torii *et al*., 1996). Therefore, L*er* was compared to its wild

relative L*ER* at different population densities. Due to the short, compact stature of L*er*, it was also hypothesised that this mutant would have an improved HI ratio compared to L*ER*.

Figure 6.1: Geographical distribution of *Arabidopsis thaliana* ecotypes used in this study.
Bla = Blanes/Gerona; Col = Columbia; Cvi = Cape Verde Islands; Kas = Kashmir; L*ER* = Landsberg *ERECTA*; Ob = Oberursel/Hasen; Rsch = Rschew/Starize.

This chapter also describes experiments that investigated the plasticity in seed size in *Arabidopsis* in response to source-sink alterations. There are many examples in *Arabidopsis* where natural or artificial reduction in seed number is accompanied by an increase in seed size (Sills and Nienhuis, 1995; Alonso-Blanco *et al*., 1999; Jofuku *et al*., 2005; Schruff *et al*., 2006). In the present study restricted pollinations, where seed set was restricted to 6 siliques on the primary inflorescence, resulted in a dramatic increase in seed size compared to unrestricted pollinations (Section 3.2.3). To gain a greater understanding of the mechanisms that enable plants to deliver heavier seeds in response to such a reduction in seed number, seed growth and development were studied in detail following restricted and unrestricted pollinations.

6.2 Results

6.2.1 Analysis of seed size, yield and HI in Arabidopsis

In order to investigate the relationship between yield and its components in *Arabidopsis*, the seed size, seed yield and HI of several ecotypes was determined. The *Arabidopsis* ecotypes chosen for this study were selected to represent the range of seed sizes found in natural populations. Rsch-0 and Ob-0 were reported to have relatively small seeds whereas Bla-1 had large seeds (Krannitz *et al*., 1991). Both Cvi (Alonso-Blanco *et al*., 1999) and Kas-1 (Ungru *et al*., 2008) have also been shown to have large seeds. Additionally, the most commonly used laboratory strains Col-0 and L*er* were analysed and, as the *erecta* mutation was hypothesised to confer an advantage due to its compact stature, both Col-*er* and L*ER* were also included in the analysis.

Despite selecting the *Arabidopsis* ecotypes for a range of seed sizes, under the highly controlled conditions used in the present study, most ecotypes had a mean seed weight of approximately 20 µg (Figure 6.2, Figure 6.3a). In this investigation, the mean seed weights of Rsch-0, Ob-0 and Bla-1 were 19.22 µg, 20.74 µg and 20.18 µg respectively, conflicting with previously documented data (Krannitz *et al*., 1991). However, Cvi and Kas-1 maintained their reportedly high seed weight, weighing 31.78 µg and 26.30 µg respectively. The nine ecotypes tested showed a wide variation in seed yield, from 0.79 g/plant in Bla-1 to 1.34 g/plant in L*ER* (Figure 6.3b). L*ER* and Rsch-0 produced the highest seed yields, suggesting a positive correlation between plant biomass and seed yield. L*ER* was particularly tall in comparison to the other ecotypes (Figure 6.4), while Rsch-0 had extremely large rosette leaves (Figure 6.5). The ecotypes also showed a wide range in HI ratios (Figure 6.3c), with the highest HI ratios found in Col-*er*, L*er* and Kas-1. These three ecotypes all have a relatively small stature (Figure 6.4).

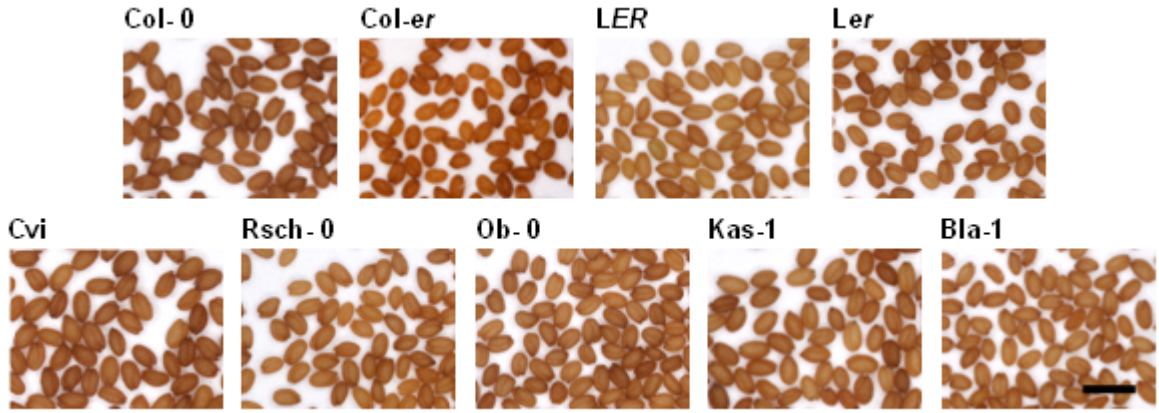

Figure 6.2: Seed phenotype of *Arabidopsis* ecotypes.

Photographs of mature dry seeds following unrestricted pollinations. Bar = 1 mm.

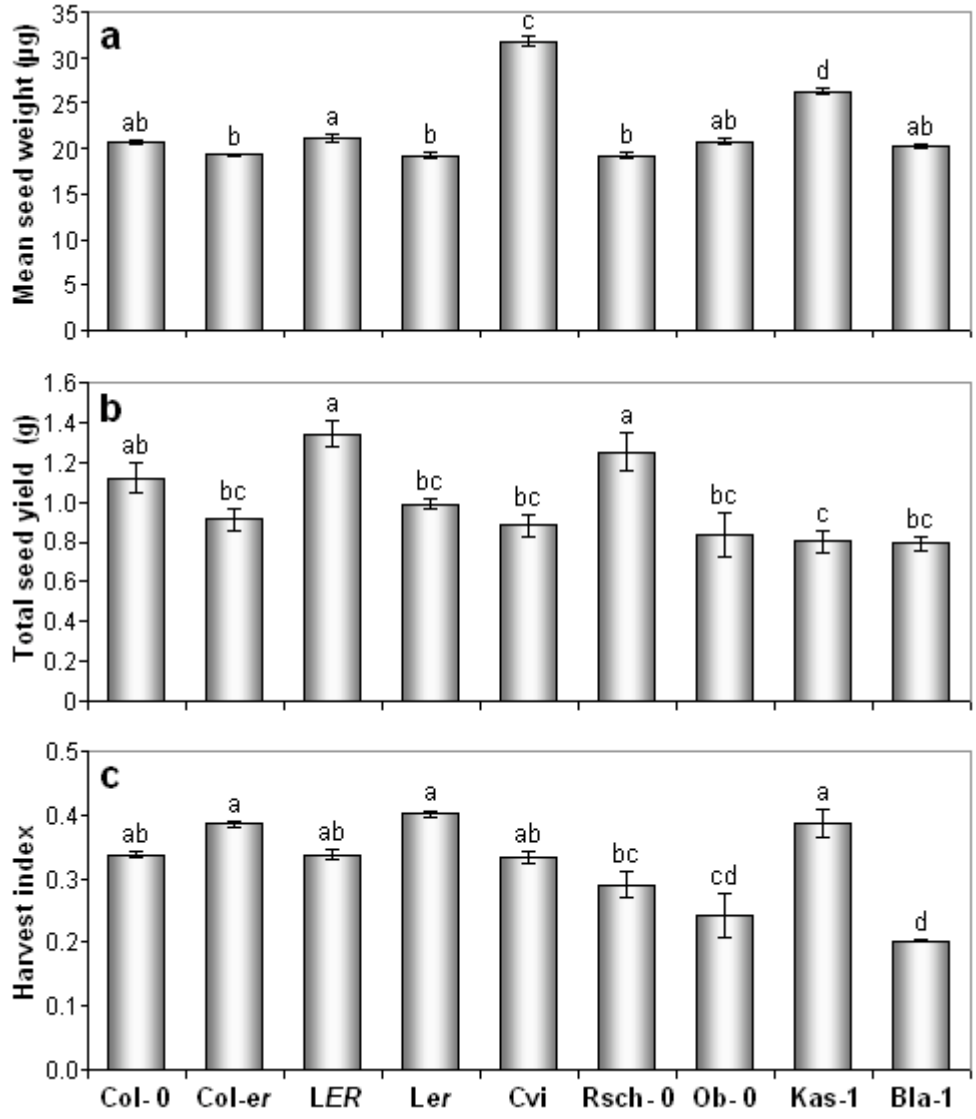

Figure 6.3: Variation in seed size, yield and HI in *Arabidopsis thaliana*.

Nine *Arabidopsis* ecotypes were assessed for (a) mean seed weight in unrestricted pollinations, (b) total seed yield per plant, and (c) HI. Values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. Error bars = s.e.m. n ≥ 6 plants.

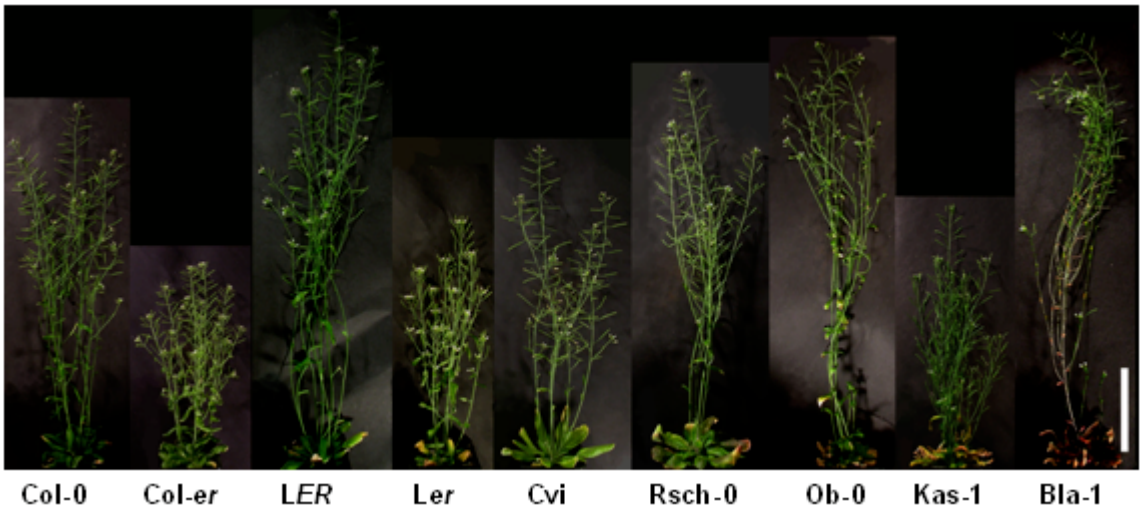

Figure 6.4: Plant stature varies greatly among *Arabidopsis* ecotypes.

Photographs of whole plants 15 days after flowering. Bar = 10 cm.

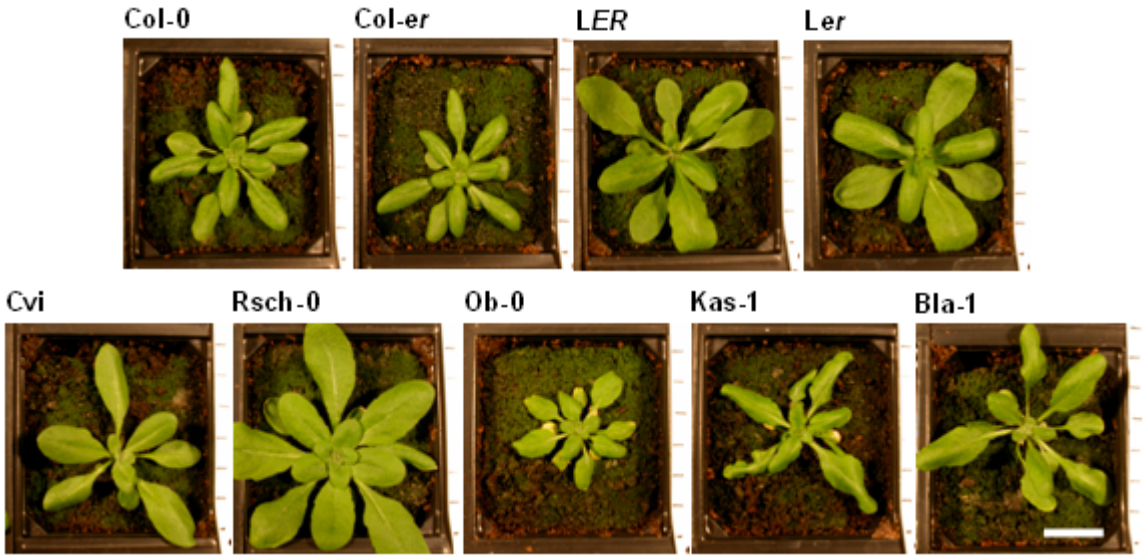

Figure 6.5: Rosette morphology varies greatly among *Arabidopsis* ecotypes.

Photographs of rosettes at 25 days after germination. Bar = 2 cm.

6.2.2 Relationships between seed yield components

The seed weight, yield and HI data from the nine *Arabidopsis* ecotypes was used to test for any significant relationships between seed yield and yield components (seed number and seed size). Seed number per plant appeared highly correlated with seed yield (Pearson's correlation, r = 0.894, P = 0.001; Figure 6.6a). However, seed weight was not associated with seed yield (Figure 6.6b); *Arabidopsis* plants that produced large seeds did not always have high yield. For example, Cvi seeds were 54.0% heavier than those of Col-0, yet the seed yield of Cvi was 21.4% lower. Seed weight was negatively correlated with seed number (Pearson's correlation, r = -0.715, P = 0.031; Figure 6.6c), showing that a trade-off exists between seed number and seed size in *Arabidopsis*.

Other factors that could affect yield include ovule number per silique, flowering time and the extent of vegetative growth, which is represented by the dry mass of aerial organs (Table 6.1). Ovule number per silique was positively correlated with seed yield (Pearson's correlation, r = 0.668, P = 0.049, Figure 6.6d). Ovule number appeared to influence yield through the total seed number per plant, as the correlation between ovule number per silique and seed number (Pearson's correlation, r = 0.601, P = 0.087) was much stronger than between ovule number per silique and seed weight (Pearson's correlation, r = -0.266, P = 0.489). In contrast, flowering time and the dry mass of aerial organs were not correlated with seed yield. Therefore, plants with a large vegetative mass did not always generate high seed yield.

Seed weight, seed number and seed yield were not associated with HI. However, HI was strongly correlated with the dry mass of aerial organs (Pearson's correlation, r = -0.859, P = 0.003, Figure 6.6e). This indicates that plants with short stature have an improved HI ratio, for example the *erecta* mutation present in both Col-*er* and L*er* reduced plant stature compared to Col-0 and L*ER* (Figure 6.4), and as a result the HI was improved (Figure 6.2c). Flowering time was also negatively correlated with HI (Pearson's correlation, r = -0.688, P = 0.040, Figure 6.6f), probably because later flowering plants undergo further vegetative growth prior to flowering.

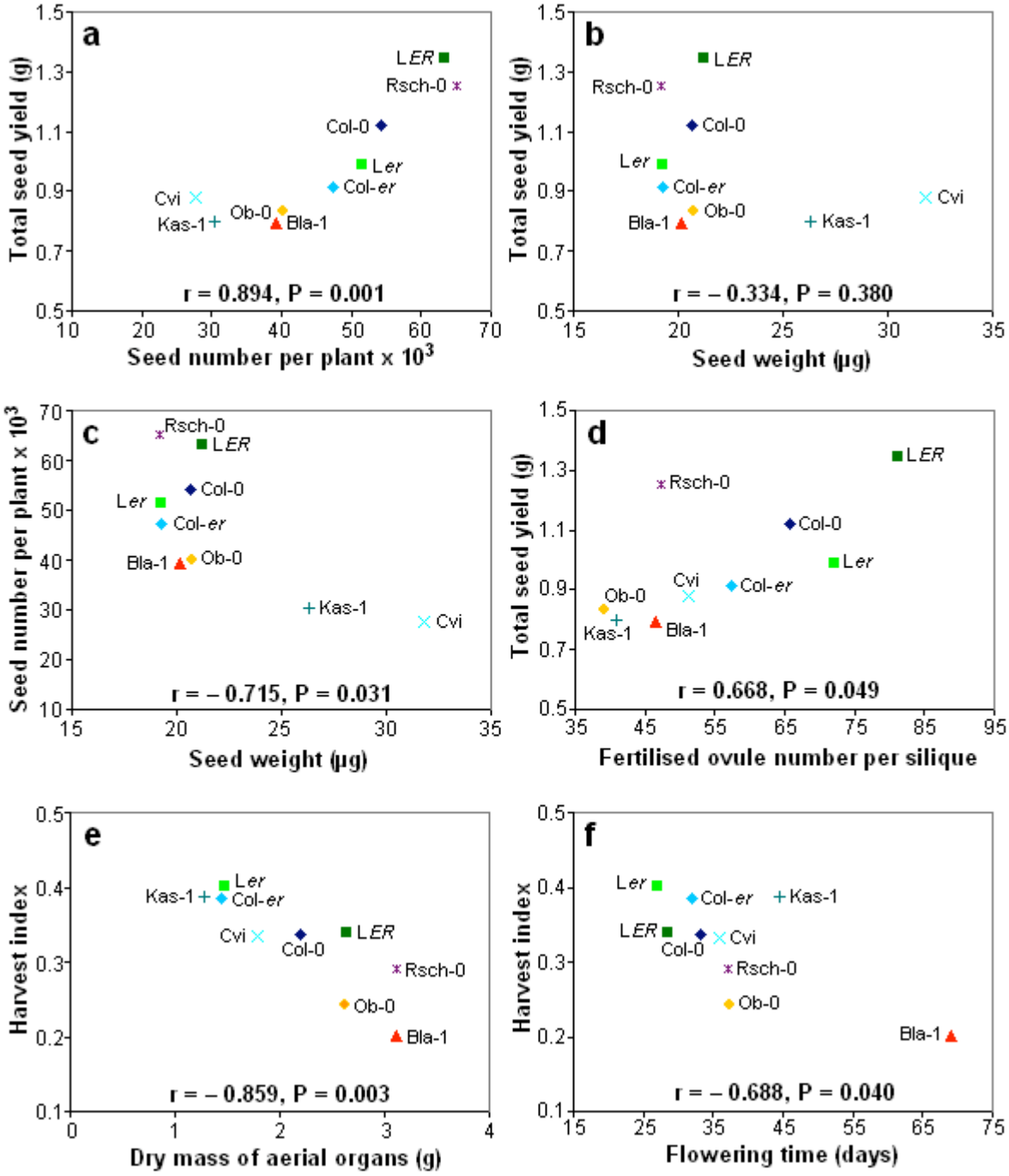

Figure 6.6: Relationship between yield and its components in *Arabidopsis*.

Correlation between seed number per plant and total seed yield (a), mean seed weight and total seed yield (b), mean seed weight and seed number per plant (c), fertilized ovule number per silique and total seed yield (d), dry mass of aerial organs and HI (e) and flowering time and HI (f). Relationships were analysed using Pearson's correlation coefficient (r).

Table 6.1: Other factors that can affect seed yield

Ecotype	Flowering time (days)	Number of fertilized ovules per silique	Flower number per plant	Dry mass of aerial organs (g)
Col-0	33.3 ± 0.25	65.8 ± 3.01	ND	2.20 ± 0.15
Col-*er*	32.0 ± 0.00	57.4 ± 0.60	ND	1.45 ± 0.09
L-*ER*	28.5 ± 0.87	81.3 ± 2.36	660.8 ± 23.9	2.63 ± 0.14
L-*er*	27.0 ± 0.00	72.2 ± 2.33	561.2 ± 21.2	1.48 ± 0.07
Cvi	36.0 ± 0.84	51.2 ± 1.53	ND	1.79 ± 0.16
Rsch-0	37.2 ± 1.07	47.4 ± 13.90	ND	3.11 ± 0.27
Ob-0	37.3 ± 1.31	39.3 ± 5.11	ND	2.63 ± 0.18
Kas-1	44.6 ± 2.64	40.8 ± 1.02	ND	1.29 ± 0.11
Bla-1	69.0 ± 1.10	46.6 ± 2.64	ND	3.12 ± 0.13

Mean ± s.e.m. Flowering time scored as the period from germination to opening of the first flower. Ovule number per silique was determined in the 20th silique on the primary inflorescence. ND = not determined.

6.2.3 Density effects on seed size, yield and HI

In the experiments reported above, yield was analysed on a per plant basis. However, from an agronomic perspective, yield should be evaluated at the level of the plant community i.e. a crop population within a field. To accommodate this, an experiment was designed to assess the effect of different planting densities on seed size, yield and HI. Trays containing a specified weight of compost were used to grow plants at three densities: 375, 750 and 1,500 plants/m^2. As a compact stature may confer an advantage at high planting densities, two genotypes were analysed, L*ER* and L*er*, as they differed only at the *ERECTA* locus, which altered plant stature. Three biological replicates were performed for each genotype at each planting density.

Planting density had no significant effect on individual seed weight (ANOVA followed by Tukey's multiple comparisons, L*er*, P = 0.136; L*ER*, P = 0.835) (Figure 6.7a). In addition, the total seed yield/m^2 (ANOVA followed by Tukey's multiple comparisons, L*er*, P = 0.151; L*ER*, P = 0.929) and the dry mass of aerial organs/m^2 (ANOVA followed by Tukey's multiple comparisons, L*er*, P = 0.095; L*ER*, P = 0.452) were not affected by plant density or genotype (Figure 6.7b,c). Thus, changes in plant density did not lead to changes in total biomass/m^2 (ANOVA followed by Tukey's multiple comparisons, L*er*, P = 0.140; L*ER*, P = 0.647) or HI (ANOVA followed by Tukey's multiple comparisons, L*er*, P = 0.119; L*ER*, P = 0.758) (Figure 6.7d,e). However, seed yield per plant changed dramatically with altered density (Figure 6.7f). By doubling the planting density from 375 to 750 plants/m^2, or from 750 to 1500 plants/m^2, the seed yield per plant was reduced by half. Since seed size remained constant, differences in seed number per plant must account for the change in seed yield per plant.

When assessed as individual plants grown under optimal conditions, the HI of L*er* (0.402 ± 0.005) was significantly higher than L*ER* (0.338 ± 0.008) (Figure 6.2c). However, when grown as a population, the HI of the *erecta* mutants was not significantly different from its wild relative at any of the densities tested (Figure 6.7e). In principal, the small stature of *erecta* mutants would require fewer resources for vegetative development, allowing more resources for reproductive development. However, L*er* mutants produced significantly lower seed yields than L*ER* plants at all planting densities (Student's *t*-test, 375 plants/m^2, P = 0.005; 750 plants/m^2, P = 0.004; 1,500 plants/m^2, P = 0.029). It is possible that L*er* plants fail to achieve high yields as they produce fewer flowers per plant (L*er*, 561.2 ± 21.2; L*ER*, 660.8 ± 23.9), fewer seeds per silique (L*er*, 72.2 ± 2.3; L*ER*, 21.2 ± 2.4, Table 6.1) and smaller seeds (L*er*, 19.2 ± 0.3 µg; L*ER*, 21.2 ± 0.4 µg, Figure 6.2a) than L*ER* plants.

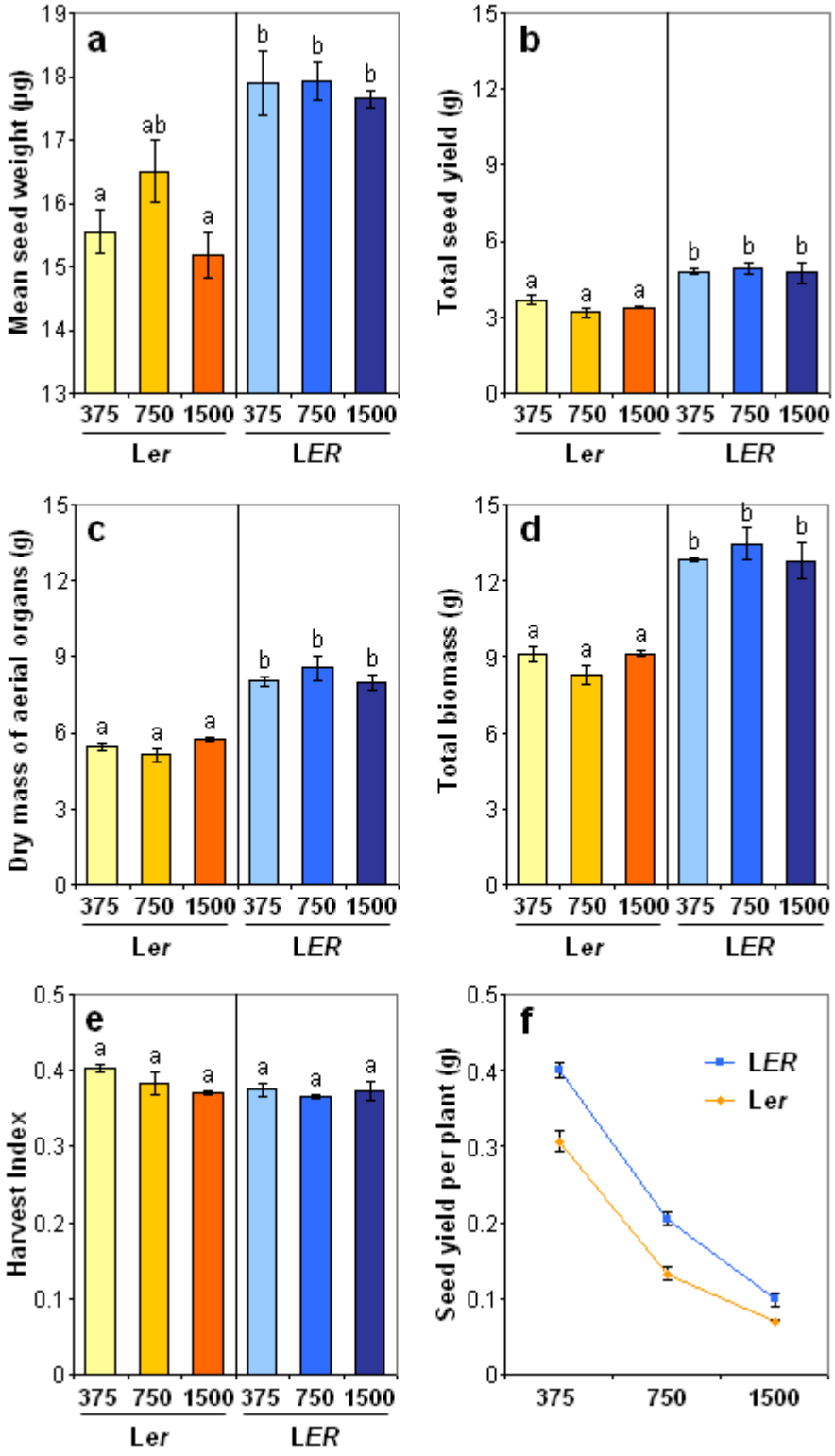

Figure 6.7: Density effects on seed size, yield and HI in L*er* and its wild relative L*ER*.
Comparison of mean seed weight (a), total seed yield (b), dry mass of aerial organs (c), total biomass (d), HI (e) and seed yield per plant (f) of L*er* and L*ER* at three planting densities: 375, 750 and 1,500 plants/m^2. For each graph, values that differ at the 0.05 significance level (ANOVA followed by Tukey's multiple comparisons) are labelled with different letters. Error bars = s.e.m. n = 3.

6.2.4 Comparison of seed growth in restricted and unrestricted pollinations

To investigate the physiological and developmental mechanisms underlying the plasticity of seed size in *Arabidopsis* in response to source-sink alterations, final seed size was compared in restricted and unrestricted pollinations (Section 3.2.3). To artificially reduce seed number, *Arabidopsis* plants were pruned to allow only six siliques to develop. The seeds harvested from these siliques were significantly heavier than seeds from unrestricted pollinations (Figure 6.8). This was demonstrated in five ecotypes: Col-0, Col-*er*, L*ER*, L*er* and Cvi (Student's *t*-test, restricted > unrestricted seed weight: for all ecotypes, $P < 0.001$). There was a considerable difference in the percentage uplift in seed weight between ecotypes, ranging from 35.0% in Cvi to 66.2% in Col-0. Interestingly, despite the high seed weight of Cvi in unrestricted pollinations, Cvi nevertheless showed substantial uplift in seed weight in restricted pollinations.

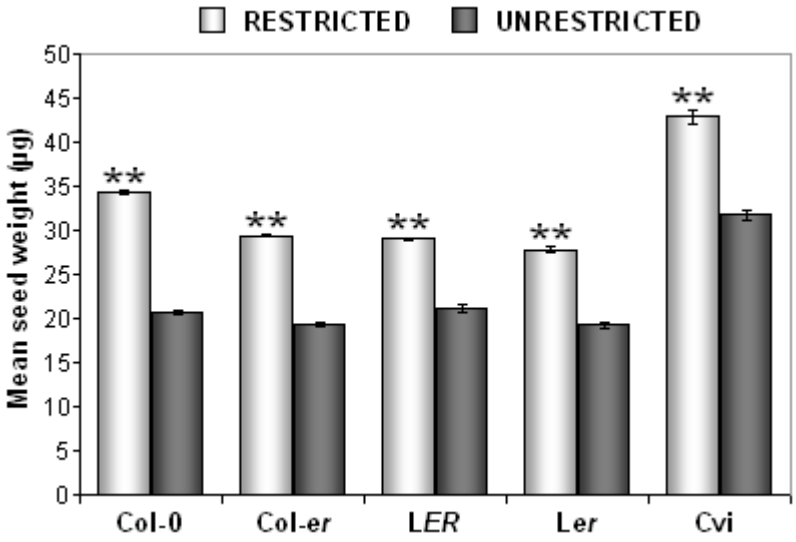

Figure 6.8: Effect of artificially restricting seed set on final seed weight.

Comparison of seed weight in restricted and unrestricted pollinations in Col-0, Col-*er*, L*ER*, L*er* and Cvi. Seed weight values that differ at the 0.01 significance level (Student's *t*-test, restricted > unrestricted) are labelled with **. Error bars = s.e.m. $n \geq 6$ plants.

The increase in seed weight in restricted pollinations compared to unrestricted pollinations demonstrates that seed size can be an extremely plastic trait in *Arabidopsis*. In order to determine how the uplift in seed weight was achieved, a detailed study of the growth of developing seeds in restricted and unrestricted pollinations was carried out in Col-0, L*er* and Cvi. The fresh weight and dry weight of developing seeds was measured every two days from 2 to 18 DAP, and at maturity (silique shattering day). The percentage water content could subsequently be calculated by subtracting the dry weight from the fresh weight.

In general, seed growth in all three ecotypes, in both restricted and unrestricted pollinations, followed a similar pattern (Figure 6.9). The fresh weight of developing seeds increased dramatically between 2 and 6 DAP. Following this period of rapid growth, the fresh weight continued to rise steadily until approximately 14 to 16 DAP, after which a sharp decrease in fresh weight was recorded. In contrast, the accumulation of dry weight in developing seeds followed a sigmoid pattern, with a lag phase from 2 to 6 DAP, a rapid phase from 6 to 14 DAP, and a final phase when little further weight was gained. The percentage water content decreased almost linearly from 80-90% in 2 DAP seeds to 5-10% in mature seeds.

By monitoring seed weight throughout seed development, it was possible to determine the time point at which an increase in seed weight was first detected in restricted compared to unrestricted pollinations. In Col-0, a significant difference in the fresh weight of seeds in restricted and unrestricted pollinations was first detected at 2 DAP (Student's *t*-test, $P < 0.001$); whereas a difference in dry weight was first apparent at 4 DAP (Student's *t*-test, $P = 0.012$). In L*er*, a significant difference in both fresh weight (Student's *t*-test, $P < 0.001$) and dry weight (Student's *t*-test, $P = 0.025$) was first observed at 4 DAP. In Cvi, a significant difference was first observed in fresh weight at 2 DAP (Student's *t*-test, $P = 0.007$) and in dry weight at 8 DAP (Student's *t*-test, $P < 0.001$).

As final seed size is a product of its two components, SGR and SFD (Egli, 1998), the uplift in seed size observed in restricted pollinations must be a consequence of either increased rate or a longer period of growth, or a combination of both. The SGR was determined during the rapid phase of dry matter accumulation between 6 and 14 DAP, and was higher in restricted pollinations than in unrestricted pollinations in Col-0 (restricted = 3.09 μg seed^{-1} day^{-1}, unrestricted = 1.34 μg seed^{-1} day^{-1}), L*er* (restricted = 2.18 μg seed^{-1} day^{-1}, unrestricted = 1.66 μg seed^{-1} day^{-1}) and Cvi (restricted = 3.96 μg seed^{-1} day^{-1}, unrestricted = 2.89 μg seed^{-1} day^{-1}). The SFD was defined as the period between fertilisation and physiological maturity (maximum dry weight). In restricted pollinations, SFD in Col-0 was 20 days, while in unrestricted pollinations it was 16 days. However, in

L*er* and Cvi the SFD for restricted and unrestricted pollinations were the same; both required 18 days to reach maximum dry weight. As an alternative method of calculating the growth period of seed development, the number of days from pollination to silique shattering was determined. In both restricted and unrestricted pollinations, Col-0 siliques shattered at 20 DAP. In restricted pollinations L*er* siliques shattered at 19 DAP whereas in unrestricted pollinations they shattered at 20 DAP. In contrast, Cvi siliques from restricted pollinations took 4 days longer (23 DAP) to shatter than those from unrestricted pollinations (19 DAP).

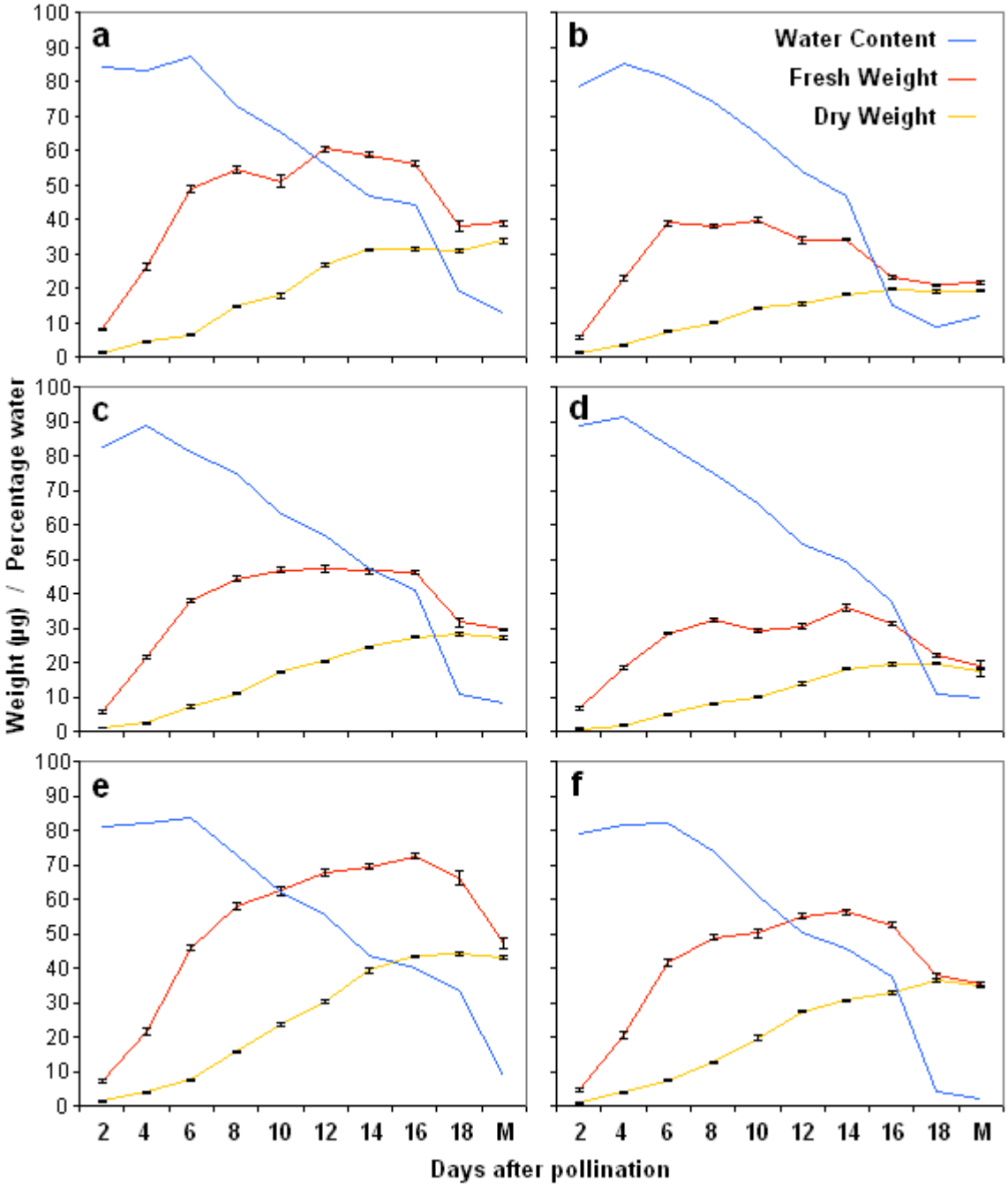

Figure 6.9: Comparison of growth of developing *Arabidopsis* seeds in restricted and unrestricted pollinations.

Changes in fresh weight, dry weight and percentage water content of developing seeds in (a) Col-0 restricted, (b) Col-0 unrestricted, (c) L*er* restricted, (d) L*er* unrestricted, (e) Cvi restricted, (f) Cvi unrestricted pollinations. M = mature. Error bars = s.e.m. n ≥ 10 batches of 10 seeds.

The relationships between final seed mass and rate and duration of seed growth were compared in order to determine which factor was associated with the increase in seed size in restricted pollinations (Figure 6.10). Final seed weight was positively correlated with SGR (Pearson's correlation, r = 0.979, P = 0.001). However, there was no correlation between the weight of mature seeds and the SFD, or the weight of mature seeds and the number of days from pollination to silique shattering. Therefore, the uplift in seed weight that occurred when seed set was restricted was associated with an increase in the rate of seed growth rather than an increase in the growth duration.

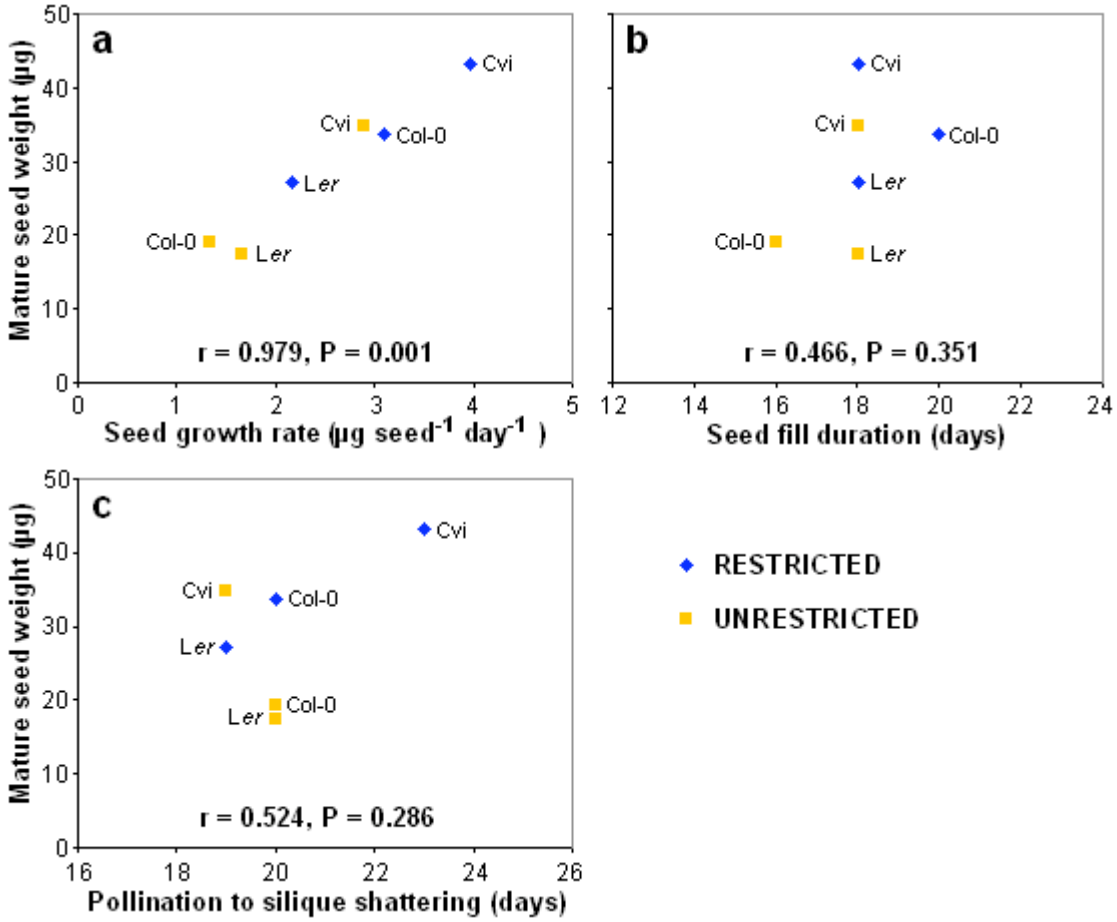

Figure 6.10: Relationship between seed weight and the rate and duration of seed growth. Correlation between mature seed weight and SGR (a), SFD (b) and number of days from pollination to silique shattering (c). Relationships were analysed using Pearson's correlation coefficient (r).

6.2.5 Determining the physiological and developmental basis for increased seed size in restricted pollinations

The uplift in seed weight observed in restricted pollinations appeared to initiate early in seed development. This was first observed for fresh weights between 2 and 4 DAP and for dry weights between 4 and 8 DAP (Figure 6.9). Seeds generated under restricted pollinations were heavier presumably due to a greater accumulation of reserves during maturation. As reported above in *Arabidopsis*, increased seed weight was associated with an increased SGR rather than SFD (Figure 6.10). The following experiments aim to determine whether the uplift in seed weight in restricted pollinations was facilitated by a larger endosperm and/or seed coat, or whether the availability of more resources simply stretched a normal-sized seed. In order to determine the physiological and developmental basis for increased seed size in restricted pollinations, endosperm and seed coat development were assessed in Col-0, L*er* and Cvi seeds from restricted and unrestricted pollinations.

Development of the integuments and subsequently the seed coat has considerable influence on seed size (Egli, 1990; Weber *et al*., 1996; Garcia *et al*., 2005). For instance, in *arf2* mutants increased cell division in the integuments generates an enlarged seed coat which is associated with the production of a heavier seed (Schruff *et al*., 2006). It is therefore reasonable to assume that the large seeds produced in restricted pollinations could have altered seed coat development. Changes in seed coat development were assessed by measuring the number of cells in three seed coat layers, ii1, oi1 and oi2 at 5 DAP (Figure 6.11). These layers were chosen for this analysis as they surround the entire embryo sac and the individual cells are clearly visible facilitating cell counting. In Col-0, a significant difference in the number of seed coat cells between restricted and unrestricted pollinations was observed in only the oi1 layer, where a 7.4% increase was seen (Student's *t*-test, oi1, $P < 0.001$). In L*er*, a significant difference was observed in both layers of the outer integument (Student's *t*-test, oi1, $P = 0.016$; oi2, $P = 0.001$); the oi1 and oi2 layers contained 5.8% and 5.9% more cells in restricted pollinations respectively. In Cvi, all layers of the seed coat tested contained more cells in restricted pollinations than in unrestricted pollinations (Student's *t*-test, ii1, $P = 0.002$; oi1, $P = 0.004$; oi2, $P = 0.005$); the ii1, oi1 and oi2 layers contained 8.1%, 5.5% and 5.4% more cells respectively. These results show a modest increase in the number of seed coat cells when seed set is restricted; however the increase in the number of integument cells appears small relative to the magnitude of the seed size uplift, especially in Col-0. However, in general Cvi seeds contained more seed coat cells than Col-0, and Col-0 contained more than L*er*, which correlated with their final seed weight.

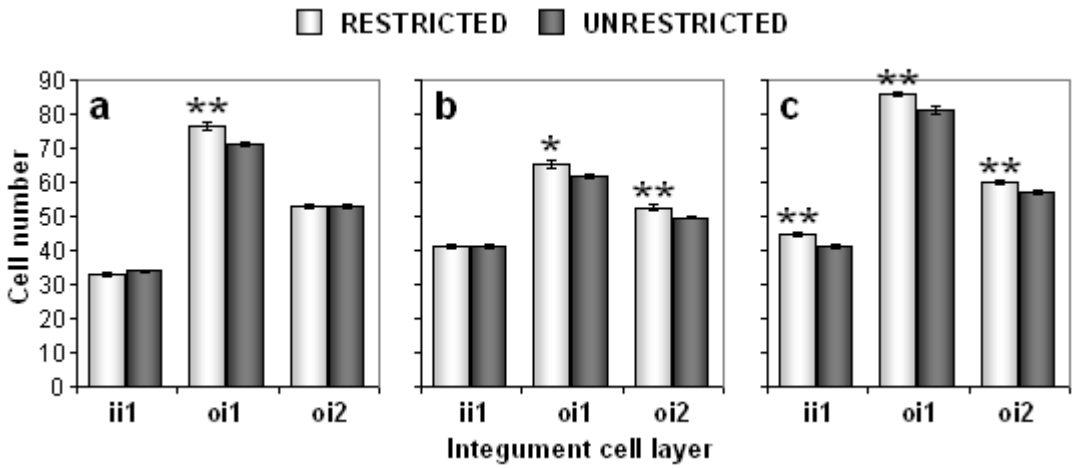

Figure 6.11: Seed coat development in restricted versus unrestricted pollinations.
Comparison of cell numbers in layers ii1, oi1 and oi2 of the seed coat at 5 DAP in Col-0 (a), L*er* (b) and Cvi (c) seeds in restricted and unrestricted pollinations. Values that differ at the 0.05 significance level (Student's *t*-test, restricted vs. unrestricted) are labelled with *, and at the 0.01 significance level with **. Error bars = s.e.m. n = 18.

Since large seeds often have a higher endosperm cell number than small seeds, the extent of endosperm proliferation is believed to be a major determinant of seed size (Cochrane and Duffus, 1983; Scott *et al*., 1998). Interploidy crosses between a diploid seed parent and a tetraploid pollen parent result in large seeds which display several characteristic features of endosperm over-proliferation, including a greater number of endosperm nuclei, delayed cellularisation and an enlarged chalazal endosperm (Scott *et al*., 1998). Thus, it is possible that the large seeds produced in restricted pollinations undergo endosperm over-proliferation in order to facilitate increased nutrient uptake and contribute to increased final seed size. To establish whether seeds produced in restricted pollinations contained over-proliferated endosperms compared to seeds from unrestricted pollinations, several aspects of endosperm development were studied: embryo sac area, chalazal endosperm area, endosperm nuclei number, endosperm cell size and the timing of endosperm cellularisation.

The area of the embryo sac (Section 3.2.3) and the chalazal endosperm were measured in 5 DAP seeds (Figure 6.12). The embryo sac area was significantly greater at this time point in all three ecotypes in restricted pollinations compared to unrestricted pollinations (Student's *t*-test, Col-0, $P = 0.002$; L*er*, $P = 0.001$; Cvi, $P < 0.001$). Embryo sac area also correlates with final seed size among ecotypes, as the embryo sac area and the final seed size of Cvi was larger than Col-0, which in turn was larger than L*er* (Figure 6.8). In contrast, the area of the chalazal endosperm was not significantly different between

restricted and unrestricted pollinations in any of the ecotypes tested. However, a difference in chalazal endosperm size was found between the ecotypes in unrestricted pollinations. Although L*er* seeds are smaller than Col-0, the chalazal endosperm area was significantly larger (Student's *t*-test, Col-0 vs. L*er*, $P < 0.001$). The chalazal endosperm of Cvi was found to be significantly larger than L*er* (Student's *t*-test, Cvi vs. L*er*, $P = 0.002$).

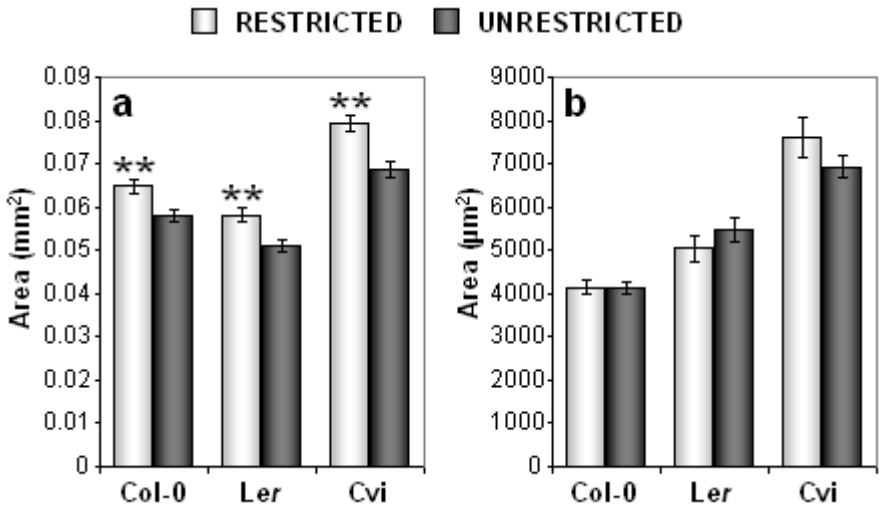

Figure 6.12: Embryo sac and chalazal endosperm area in restricted and unrestricted pollinations.

Comparison of the embryo sac (a) and chalazal endosperm (b) area at 5 DAP in Col-0, L*er* and Cvi developing seeds in restricted and unrestricted pollinations. Values that differ at the 0.01 significance level (Student's *t*-test, restricted vs. unrestricted) are labelled with **. Error bars = s.e.m. (a) $n \geq 39$. (b) $n \geq 7$.

The results presented in Figure 6.12 show that developing seeds from restricted pollinations have a significantly larger embryo sac area than seeds from unrestricted pollinations. Therefore, more space is available for endosperm growth in these seeds. To determine whether endosperm growth was increased, the number of free endosperm nuclei was counted in seeds containing early heart stage embryos (approximately 4 DAP) in restricted and unrestricted pollinations. Feulgen stained seeds were visualised using confocal microscopy. Only seeds orientated so that the embryo and the chalazal endosperm cyst were equally visible were used. Depending on seed size, 10-17 layers were viewed in order to count the number of endosperm nuclei present in half a seed. This result was doubled in order to give the total nuclei number per seed. Care was taken to avoid counting the same nuclei twice. Despite the presence of more nuclei in all three ecotypes in

restricted pollinations compared to unrestricted pollinations (Col-0, 32.4%; L*er*, 13.0%; Cvi, 5.2%), no statistically robust difference was found (Figure 6.13). However, the number of endosperm nuclei did vary between the ecotypes, with Cvi containing the highest number of nuclei at this time point.

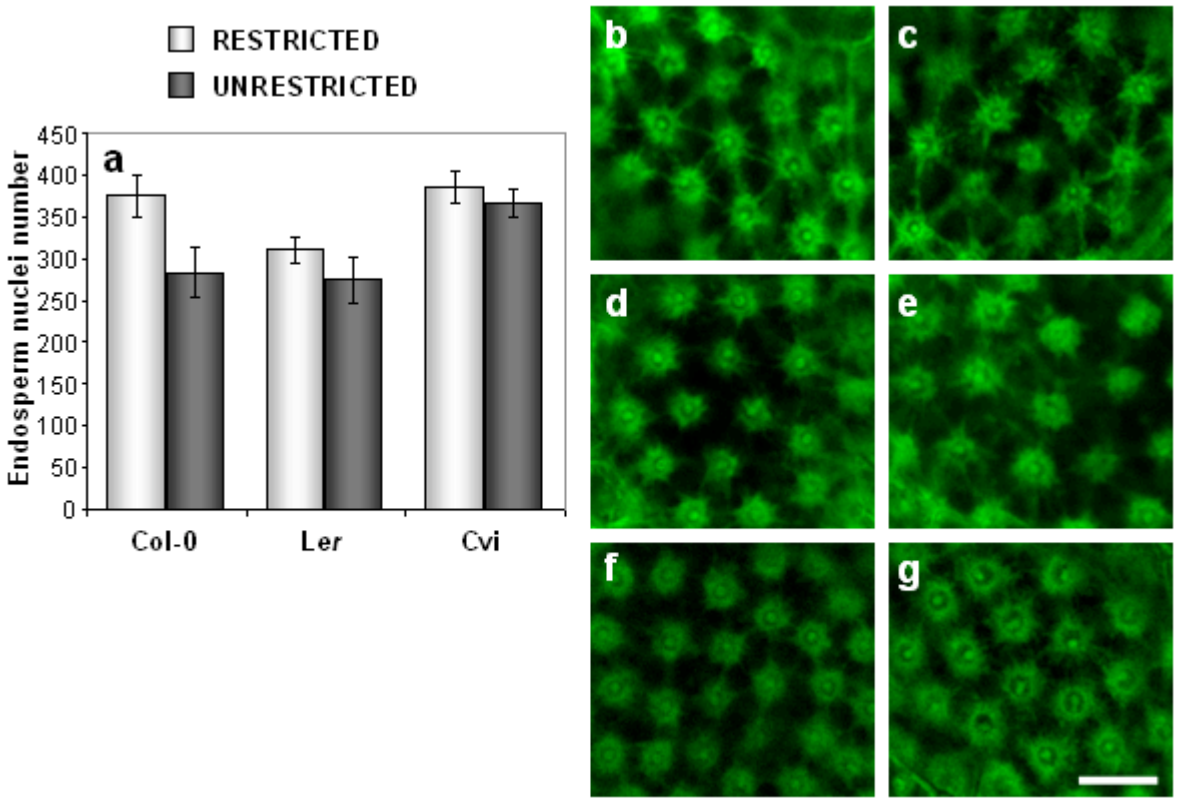

Figure 6.13: Endosperm proliferation in restricted versus unrestricted pollinations.
(a) Comparison of endosperm nuclei number in Col-0, L*er* and Cvi developing seeds containing early heart stage embryos (approx. 4 DAP) in restricted and unrestricted pollinations. n = 5. Error bars = s.e.m. Density of free endosperm nuclei in Col-0 restricted (b), Col-0 unrestricted (c), L*er* restricted (d), L*er* unrestricted (e), Cvi restricted (f) and Cvi unrestricted (g). Bar = 50 µm.

In order to provide further evidence that increased endosperm proliferation occurs in restricted pollinations, the area of endosperm cells in the outermost cellularised layer (aleurone) was measured in seeds containing torpedo stage embryos (approximately 6 DAP). The data showed that there was no significant difference between the area of aleurone cells in restricted versus unrestricted pollinations or between Col-0 and Cvi (Figure 6.14), suggesting that endosperm cell size is highly stable. If the cells in the endosperm of seeds from restricted and unrestricted pollinations are of the same size, the larger cavity in restricted pollinations must be filled with more cells. This suggests that

more endosperm nuclei are present prior to endosperm cellularisation in restricted pollinations. Although not statistically significant, an increase in free endosperm nuclei was observed in seeds from restricted pollinations (Figure 6.13), which may have been sufficient to fill the enlarged embryo sac.

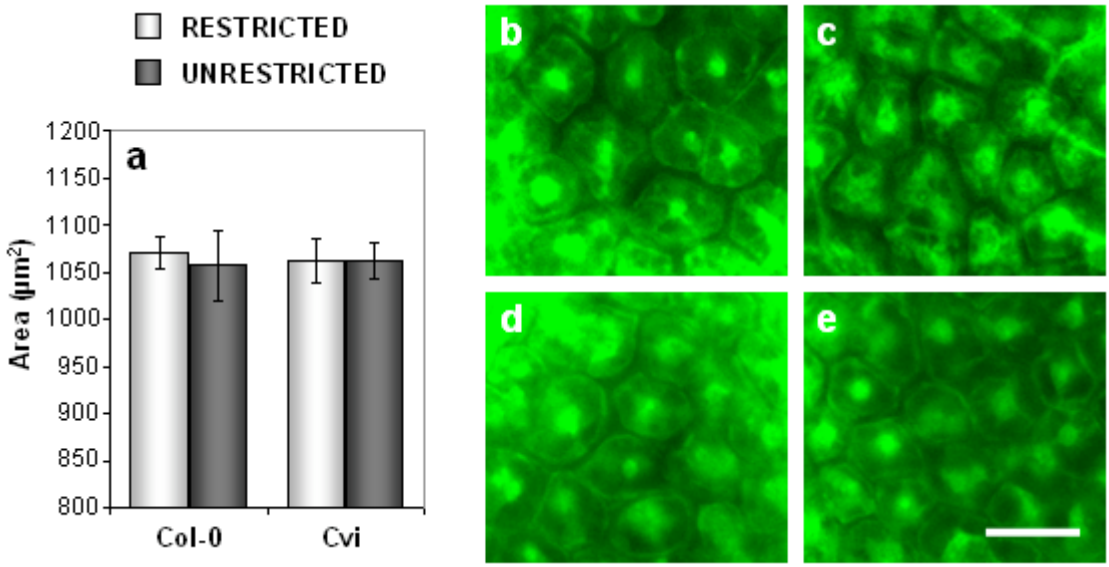

Figure 6.14: Endosperm cell size in restricted versus unrestricted pollinations.
(a) Comparison of endosperm cell area in Col-0 and Cvi developing seeds containing torpedo stage embryos (approx. 6 DAP) in restricted and unrestricted pollinations. n ≥ 90. Error bars = s.e.m. Representative images of endosperm cells in the outer endosperm layer (aleurone) in Col-0 restricted (b), Col-0 unrestricted (c), Cvi restricted (d) and Cvi unrestricted (e) seeds. Bar = 50 µm.

The timing of endosperm cellularisation can be an important factor in determining the extent of endosperm proliferation (Scott *et al*., 1998; Kang *et al*., 2008). Therefore, seeds from restricted and unrestricted pollinations were compared to find out if restricting the seed set had altered the timing of cellularisation of the endosperm. Endosperms were scored as either: uncellularized; in early stages of cellularisation (cellular endosperm surrounding the embryo and the periphery of the embryo sac); or in late stages of cellularisation (majority of embryo sac is cellularised) (Figure 6.15). Results showed that seeds with globular embryos contained mostly syncytial endosperms however 5-30% had initiated the early stages of cellularisation (Figure 6.16). The percentage of seeds with globular embryos that had initiated endosperm cellularisation was not significantly different in restricted and unrestricted pollinations. In contrast, all seeds with heart stage embryos had initiated endosperm cellularisation, with 0-20% of seeds showing almost

complete cellularisation. Again, no significant difference in the percentage of seeds in the late stages of cellularisation was observed between restricted and unrestricted pollinations. In all seeds with torpedo stage embryos, whether produced in restricted or unrestricted pollinations, and regardless of ecotype, the endosperm was almost entirely cellular. In summary, these results suggest that the timing of endosperm cellularisation was not altered when seed set was restricted, and therefore was not responsible for the uplift in seed size.

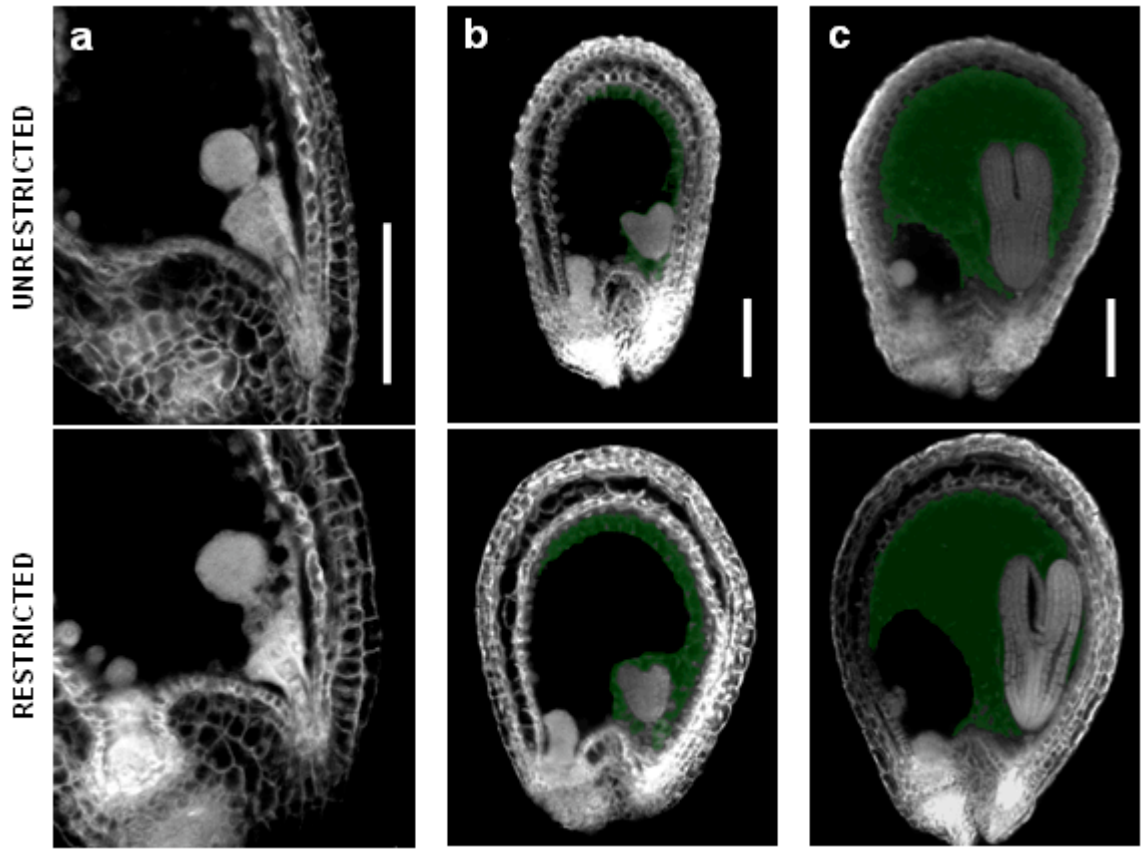

Figure 6.15: Endosperm cellularisation in restricted versus unrestricted pollinations.
Confocal images of L*er* during restricted and unrestricted pollinations showing the extent of endosperm cellularisation in developing seeds containing globular (a), heart (b) and torpedo (c) embryos. Cellularised endosperm has been false coloured in green. Bar = 100µm.

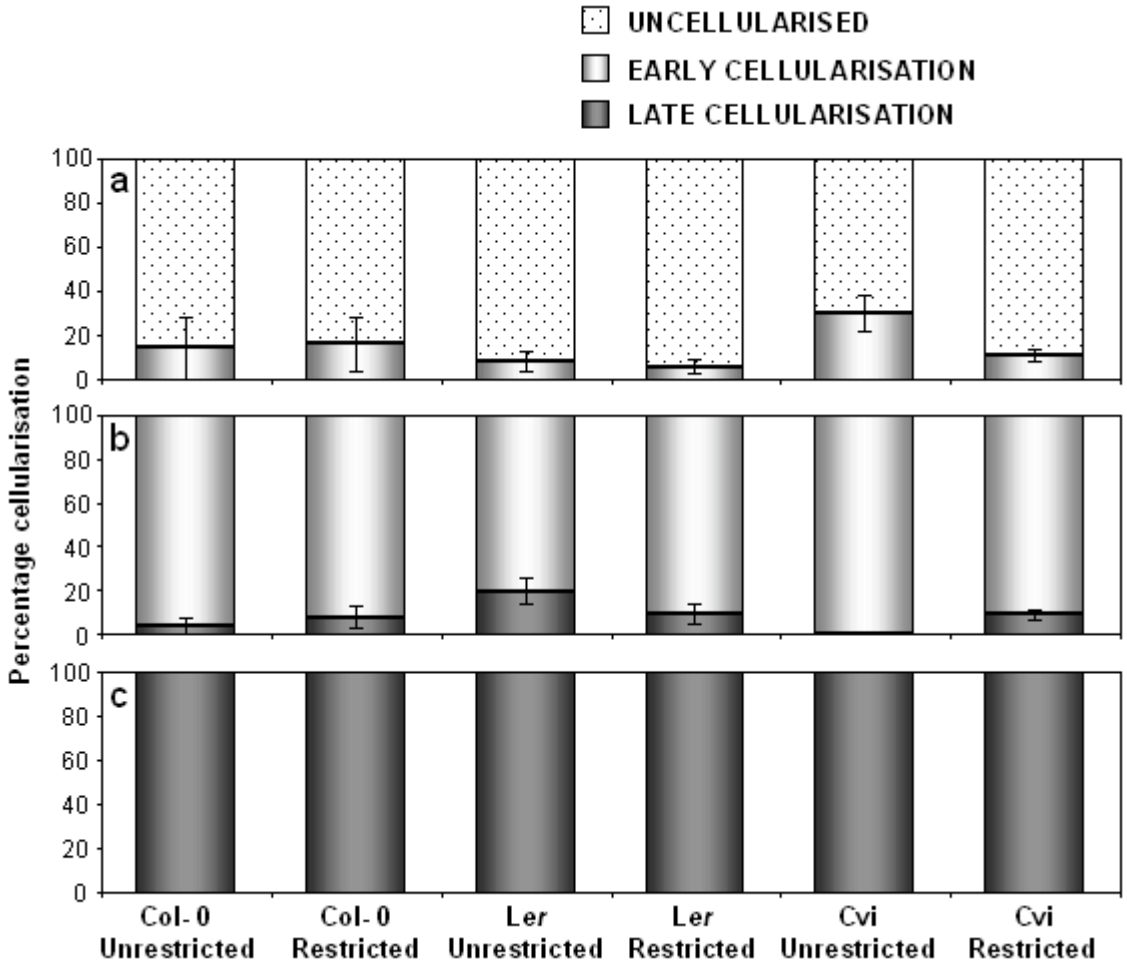

Figure 6.16: Timing of endosperm cellularisation in restricted versus unrestricted pollinations.

Comparison of the timing of endosperm cellularisation in Col, L*er* and Cvi developing. Each graph represents the extent of endosperm cellularisation at a particular embryo stage: (a) globular, (b) heart, and (c) torpedo. Error bars = s.e.m. $n \geq 38$.

6.3 Discussion

6.3.1 Seed size, yield and HI vary greatly among Arabidopsis ecotypes

The molecular networks that regulate seed yield remain largely unknown. Since *Arabidopsis* has many experimental advantages over crop species, this model plant could help elucidate important genes and their functions, and provide novel methods for crop improvement. Prior to starting an investigation into the underlying molecular mechanisms of seed yield, it was thought necessary to understand the variation in yield in *Arabidopsis* and the relationship between yield and its components. Thus, the variation in seed size, yield and HI was established for several *Arabidopsis* ecotypes (Figure 6.3).

Previous research has shown that *Arabidopsis* ecotypes exhibit a large variation in seed size (Krannitz *et al*., 1991; Alonso-Blanco *et al*., 1999; Ungru *et al*., 2008). Although the seed weights recorded in this study conflicted with existing data for some ecotypes, a wide range in seed weight was observed under our highly controlled growth conditions (Figure 6.3a). The smallest seeds were produced in the Rsch-0 ecotype and weighed 19.22 µg, while the largest seeds were found in Cvi and were approximately 65% heavier, at 31.78 µg. Before this research, the variation in yield and HI had not been explored in *Arabidopsis*. In the nine ecotypes tested yield varied by 1.7-fold, from 0.79 g/plant in Bla-1 to 1.34 g/plant in L*ER* (Figure 6.3b). HI was even more variable; with the highest ratio of 0.40 in L*er*, double that of the lowest ratio of 0.20 in Bla-1 (Figure 6.3c). The proportion of assimilates allocated to seed production therefore varies greatly among *Arabidopsis* ecotypes.

Natural variation in *Arabidopsis* has previously been utilised to identify important genetic loci that determine seed size (Alonso-Blanco *et al*., 1999). Eleven loci involved in the regulation of seed size were identified in a QTL analysis between L*er* and Cvi (Alonso-Blanco *et al*., 1999); however the genes responsible for the difference in seed size between these two ecotypes have yet to be isolated. QTL analyses have led to the cloning and characterisation of genes involved in seed size regulation in several crop species (Fan *et al*., 2006; Song *et al*., 2007; Orsi and Tanksley, 2009). Additionally, yield-enhancing QTLs have been identified in several crop species (Mallikarjuna Swamy and Sarla, 2008). Recently, the *DEP1* locus, encoding a PEBP-like domain protein, has been shown to be solely responsible for increased grain yield within a major rice grain yield QTL (Huang *et al*., 2009). This study has shown that variation exists for yield and HI in *Arabidopsis*; thus variation in *Arabidopsis* provides the opportunity for discovering further QTLs associated with yield and HI, and facilitating their functional characterisation.

6.3.2 Large seed size is not associated with high yield or high HI

The seed size, yield and HI data collected for the *Arabidopsis* ecotypes enabled the relationship between the yield components to be determined. Studying yield components is often discouraging due to yield component compensation. Indeed, this data confirmed that a trade-off between seed size and seed number exists in *Arabidopsis* (Figure 6.6c). However, analysing yield components is necessary to fully understand how yield is determined and reveal which components have the largest impact on yield. The results obtained for *Arabidopsis* concur with those previously reported for soybean and wheat (Egli, 1998), maize (Jong *et al*., 1982) and sunflower (Cantagallo *et al*., 1997). Seed number was positively correlated with yield (Figure 6.6b), whereas seed weight was not associated with yield (Figure 6.6c). Cvi which produced the heaviest seeds of all the ecotypes tested, did not have high seed yield. This ecotype produced the lowest number of seeds per plant (Figure 6.6.a) and had few ovules per silique (Table 6.1) compared to other ecotypes tested, suggesting that the low yield of Cvi is caused by low seed number (Figure 6.6.a).

The life history traits of Cvi and L*er* have been previously analysed in order to determine how increased seed size is achieved in Cvi (Alonso-Blanco *et al*., 1999). This study also found that Cvi had fewer seeds per plant and per ovule than L*er*. In addition, Cvi had a higher percentage of unfertilised ovules per silique compared to L*er*, representing one way in which reproductive effort is lost in Cvi, contributing to a reduction in yield. These results suggest that strategies to increase yield should target seed number due to the strong positive relationship between seed number per plant and yield, rather than seed weight, which is not associated with yield. One approach to increase seed number per plant is to increase the number of ovules per silique, as the number of fertilised ovules per silique was found to be positively correlated with seed yield in *Arabidopsis*.

Seed yield, seed weight and seed number were not associated with HI in *Arabidopsis*. For example, L*ER* had a high yield and a high seed number but a low HI ratio. This suggests that strategies to improve yield through either increasing seed weight or seed number will probably not result in improved HI. It was also found that flowering time was not associated with seed yield, but was negatively correlated with HI. Flowering time is an important fitness component in annual plants as a longer vegetative growth phase should allow for a greater seed output. However, delayed flowering is associated with costs that counterbalance the benefits, and subsequently each ecotype has an optimal flowering time depending on the environmental conditions in which they evolved (Kudoh *et al*., 2002). Therefore, when grown under controlled conditions it was not surprising that flowering time was not associated with yield. Late flowering ecotypes often have greater

vegetative mass due to a longer growth period but as this is not always associated with a concomitant increase in seed yield, the HI of late flowering ecotypes is low. This indicates that early flowering could be a beneficial trait in crop plants however it may not directly lead to increased seed yield.

6.3.3 ERECTA: a potential Green Revolution gene

The HI is defined as the ratio of seed yield to total biological yield (seed yield plus the dry mass of aerial organs) (Donald, 1962; Donald and Hamblin, 1976). The HI ratio is useful as it shows that yield improvements can be achieved without an increase in total plant biomass by increasing the amount of assimilates partitioned to seeds. However in *Arabidopsis*, HI was not associated with seed yield but was negatively correlated with the dry mass of aerial organs (Figure 6.6e). It was expected that short stature in *erecta* mutants would confer an improved HI ratio compared to their wild relatives. Results showed that this was true in both the L*ER* and Col-0 genetic backgrounds (Figure 6.3c). *erecta* mutants have compact inflorescences, short petioles and blunt siliques (Torii *et al*., 1996). The increased HI ratio was therefore a result of a reduction in the dry mass of aerial organs of the *erecta* mutants compared to L*ER* and Col-0. A high HI was also found in Kas-1 (Figure 6.3c), a natural accession originating from India which has small stature (Figure 6.4). Several naturally-occurring accessions have been found to carry a malfunctioning *erecta* allele, such as Hir-1 and Van-0 (Van Zanten *et al*., 2009). It would be interesting to determine whether Kas-1 is also an *erecta* mutant.

An increased HI ratio implies that fewer assimilates are used in the production of vegetative mass so that more are available for seed production. Despite the improved HI ratios of L*er* and Col-*er*, seed yield in these mutants was lower than in L*ER* and Col-0. In the case of L*er*, reduced yield compared to L*ER* was associated with the production of smaller seeds, fewer fertilised ovules per silique, and fewer flowers per plant. Although L*er* is commonly used as a wild-type, the *erecta* mutation causes more alterations to plant development than first realised (Van Zanten *et al*., 2009), some of which have the potential to negatively affect yield. In addition to mediating plant architecture, *ERECTA* influences photosynthetic capacity and the electron transport rate by affecting transpiration efficiency (Masle *et al*., 2005). This raises a key question: is there any advantage to improving HI if there is no affect on seed yield? During the Green Revolution, dwarf cereal varieties with improved HI ratios were associated with high yields due to reduced lodging compared to tall varieties when grown under high levels of nitrogen fertilizer (Khush, 1999). This suggests that yield losses were reduced but it is not clear whether the yield potential of the dwarf cereals was higher than that of the tall varieties.

Traditionally, yield has been evaluated at the level of the plant community as important factors such as plant population, leaf area index and light interception can also be considered (Egli, 1998). Therefore, the yield and HI of L*er* and L*ER* were assessed for a population of plants under different planting densities. Grain yield, total biomass and HI have been shown to vary greatly with changes in density in maize, wheat and sorghum (Donald and Hamblin, 1976). However, in *Arabidopsis* at the densities tested, no change in total seed yield, total biomass or HI was found (Figure 6.7). In this experiment, three planting densities were tested, 375, 750 and 1,500 plants/m^2, representing a 4-fold increase in the number of plants in a given area. Other studies have altered population density from 2-fold to 320-fold (Donald and Hamblin, 1976; Nelson and Ohlrogge, 1957; Begna *et al*., 1997). It is likely that the range used for *Arabidopsis* was not wide enough to affect seed yield, total biomass and HI. The findings presented here show that reduced stature in L*er* compared to L*ER* did not confer any advantage in terms of seed yield at the planting densities tested. This contrasts with previous data from maize which demonstrated that the maximum yield of a standard variety occurred at 13,000 plants/acre, and was less than the maximum yield of a dwarf variety which occurred at 26,000 plants/acre (Nelson and Ohlrogge, 1957).

Mean seed weight was not altered across the densities tested (Figure 6.7a). However, the mean seed weight values obtained during the density experiment (L*er*, 15.7 µg; L*ER*, 17.8 µg) were considerably lower than those obtained on a per plant basis (L*er*, 19.2 µg; L*ER*, 21.2 µg). It is therefore possible that competition between the plants grown as a community reduced mean seed weight compared to plants grown individually. Moreover, investigations examining the effects of population density on maize found a reduction in kernel weight with increasing stand density (Poneleit and Egli, 1979; Borrás *et al*., 2003). Whilst seed size remained constant, we found that seed yield per plant decreased dramatically with increasing plant density (Figure 6.7f). The fall in seed yield per plant must therefore have resulted from a reduction in seed number per plant. This reinforces the assumption that seed number is affected by the environment to a greater extent than seed size.

Although *erecta* mutants had reduced stature and an improved HI ratio, they produced lower yield and were not advantageous under the population densities tested. The wild-type ERECTA protein specifies aerial organ shape and size primarily through cell division (Shpak *et al*., 2003, Shpak *et al*., 2004). Therefore, *erecta* mutants are not typical dwarfs, the majority of which are a consequence of altered cell elongation. The dwarf cereal varieties that were key to the success of the Green Revolution were short due to abnormal responses to the growth-promoting hormone gibberellin. Dwarfing in the *gai*

mutant is a result of reduced cell elongation in the bolting stem (Fridborg *et al*., 1999). The brassinosteroid mutant *dwarf-4* also fails to undergo normal cell elongation (Azpiroz *et al*., 1998). Thus *ERECTA* provides an alternative mechanism of altering plant stature. Genetic manipulation of plant architecture such as this is essential for generating high-yielding crop varieties (Sakamoto and Matsuoka, 2004). Like *erecta*, the *Arabidopsis gai* and the wheat *Rht-B1b*/*Rht-D1b* homeoalleles confer phenotypes throughout the plant lifecycle and these may be detrimental to yield under certain growth conditions (Flintham *et al*., 1997; Ait-ali *et al*., 2003). However, ethanol-induced transient expression of *gai* in *Arabidopsis* enabled reduced stem growth to be achieved in isolation from other aspects of the phenotype (Ait-ali *et al*., 2003). A similar strategy could be used to exploit *ERECTA*, providing the opportunity to reduce plant stature and increase HI, without negative effects on yield. *ERECTA* therefore has the potential to become a new Green Revolution gene.

6.3.4 Seed size in Arabidopsis is extremely plastic

Source-sink alterations can cause dramatic alterations in seed size. This has been previously shown in maize (Kiniry *et al*., 1990; Borrás and Otegui, 2001; Borrás *et al*., 2003), soybean (Borrás *et al*., 2004) and sorghum (Gambin and Borrás, 2007). The present study shows this is also true for *Arabidopsis*. Comparisons of final seed mass in restricted and unrestricted pollinations in five *Arabidopsis* ecotypes showed that there was a dramatic uplift in seed size from 35.0% to 66.2% when seed set was restricted (Figure 6.8). Even Cvi, which has naturally very large seeds, undergoes considerable increase in seed size. These experiments suggest that seed size is an extremely plastic trait. However, this plasticity appears to be expressed only after seed number per plant has been set and is a consequence of seed filling being the last stage in the yield production process.

Large seed size in Cvi in unrestricted pollinations compared to other ecotypes is associated with a reduced seed number per plant (Figure 6.6a) and fewer fertilised ovules per silique (Table 6.1). However, the increased seed size must in part be due to genetic factors as in restricted pollinations, where there is a plentiful supply of resources per seed, Cvi seeds remain considerably larger than those of other ecotypes. This evidence supports the concept that growth of individual seeds is to some extent controlled by the seeds themselves (Egli, 2006). Indeed, Cvi seeds were found to contain 10% more cells in the oi2 layer of the seed coat than L*er* seeds (Alonso-Blanco *et al*., 1999). Furthermore, this study also found that Cvi seeds possess seed coats containing more cells than Col-0 and L*er* (Figure 6.11) and have a greater number of endosperm nuclei (Figure 6.13). Final seed size is therefore a product of both the 'sink capacity' of individual seeds and the availability of resources to fill the sinks.

To determine how uplift in seed size was achieved in restricted pollinations, a detailed analysis of seed growth throughout development was carried out. The general pattern of fresh weight and dry weight accumulation and percentage water loss was similar for all ecotypes under both pollination treatments (Figure 6.9). This growth pattern was almost identical to that which has been previously obtained for the Wassilewskija (Ws) ecotype (Baud *et al*., 2002). The increase in seed size was found to be associated with a faster filling rate rather than a longer filling period (Figure 6.10), in agreement with results from similar experiments in maize (Borrás and Otegui, 2001; Borrás *et al*., 2003). This suggests that developing *Arabidopsis* and maize seeds respond to source-sink alterations using an analogous mechanism.

Increasing the source-sink ratio in *Arabidopsis* increased dry matter accumulation, but also the rate of water uptake into developing seeds. The fresh weight of seeds early in development was found to be considerably greater in restricted compared to unrestricted pollinations (Figure 6.9). It is possible that this increased water uptake could contribute to increased dry matter accumulation or simply be a consequence of it. Severe water deficit in soybean plants does not affect water status or the growth rate of developing seeds, yet seed size is reduced due to shortening of the growth period (Westgate and Grant, 1989a, 1989b). Rapid cell division contributes to seed growth early in development and later growth is the result of cell expansion, which is driven by water movement into seeds. However, water uptake is driven by the osmotic gradient across the cell wall and thus assimilate availability may regulate net water uptake, cell expansion and the duration of seed growth (Egli, 1990). In maize, it has been shown that final kernel weight is closely related to maximum kernel water content (Borrás *et al*., 2003). Later in seed development, water is replaced by storage products causing a progressive desiccation which eventually limits metabolism and further growth ceases. Recently, three aquaporin-encoding genes have been identified in French bean (*Phaseolus vulgaris*) seeds that show their strongest expression in seed coats (Zhou *et al*., 2007). Phloem water is co-imported into seeds with nutrients. Aquaporins embedded in the plasma membranes of seed coat cells are responsible for water exchange to the seed apoplasm. Further work in this area is required to fully understand the role of water uptake in seed growth.

A significant difference in the weight of seeds from restricted and unrestricted pollinations was first observed remarkably early in development. This was first observed for fresh weights between 2 and 4 DAP and for dry weights between 4 and 8 DAP (Figure 6.9). As seeds respond early in development to source-sink alterations, it was possible that seed development was altered and subsequently contributed to enhanced seed growth in restricted pollinations. Dramatic changes in seed development, predominantly within the

endosperm, were observed in the large seeds produced by interploidy crosses between a diploid seed parent and a tetraploid pollen parent (Scott *et al*., 1998). The seed phenotype present is due to the presence of extra paternally-derived chromosomes in the endosperm. Seeds with paternal excess contained a greater number of endosperm nuclei than seeds from a cross between two diploid plants. Cellularisation was also delayed and the chalazal endosperm and chalazal nodules were enlarged. The effects of the interploidy cross on seed coat development have not been elucidated but the seed coat must be larger, due to increased cell division or expansion, to accommodate a larger endosperm and embryo. Therefore, potentially important aspects of seed development were studied in restricted versus unrestricted pollinations to determine the physiological basis of increased seed size in response to an increase in the source:sink ratio.

Seed coats of seeds from restricted pollinations contained a significantly larger number of cells in some layers compared to those from unrestricted pollinations in the three ecotypes tested (Figure 6.11). However, this increase in cell number does not seem great enough to account for the large increase in seed weight, especially in Col-0 in which only a single layer contained a significantly greater number of cells. This suggests that in restricted pollinations, cell expansion was also increased in the seed coat to allow additional seed growth. Cell expansion has been previously shown to compensate for reduced cell division in the seed coats of plants containing the *p35S::KRP2* transgene (Garcia *et al*., 2005). The enlarged seed coat, due to increased cell division, found in *arf2* seeds is associated with a 21% increase in seed size, even when seed set is held constant (Schruff *et al*., 2006). Importantly, the integuments of *arf2* have been shown to contain more cells than those of wild-type prior to fertilisation. It is possible that increased cell division before fertilisation is a necessary requirement for dramatic increases in cell number in the seed coat. It seems unlikely that a higher source:sink ratio would influence integument cell number prior to fertilisation. In contrast, it is possible that ecotype differences in seed coat cell number are present in the integuments prior to fertilisation, for example Cvi ovules may contain more cells in the integuments than those of Col-0 or L*er*, and this could contribute to the increased size of Cvi seeds.

Developing seeds from restricted pollinations had a significantly larger embryo sac area than seeds from unrestricted pollinations at 5 DAP (Figure 6.12a). Therefore, in restricted pollinations there was a larger seed cavity providing more space for the developing endosperm. Unlike in seeds with paternal excess, where endosperm development is altered considerably (Scott *et al*., 1998), uplift in seed size due to an increased source:sink ratio caused relatively few changes. The chalazal endosperm was not significantly different in size in restricted compared to unrestricted pollinations (Figure

6.12b). Although more endosperm nuclei were found in seeds from restricted pollinations, this difference was not significant (Figure 6.13). However, as the size of the endosperm cells was found to be extremely stable across both pollination treatments and all ecotypes, it seems likely that the increased embryo sac area must be filled with a greater number of endosperm nuclei. Nonetheless, endosperm proliferation did not appear to increase to the extent found in seeds with paternal excess (Scott *et al*., 1998). The number of endosperm nuclei was found to vary among *Arabidopsis* ecotypes and was positively correlated with final seed weight. This has also been shown in barley cultivars differing in grain weight (Cochrane and Duffus, 1983), and suggests that endosperm proliferation is an important determinant of seed size. Although there is much evidence to indicate that endosperm cellularisation has a significant effect on seed size (Scott *et al*., 1998; Kang *et al*., 2008) there was no alteration to the timing of endosperm cellularisation in restricted compared to unrestricted pollinations.

6.3.5 Further work

This chapter described an investigation into the variation in seed size, yield and HI of nine *Arabidopsis* ecotypes. More than 750 ecotypes are available from seed stock centres and so proportionally few ecotypes were tested. Sampling more ecotypes would provide a more accurate representation of the variation found in natural populations. In addition, it would improve the confidence with which relationships between yield components can be drawn. Since the *Arabidopsis* ecotypes tested exhibited a wide variation in seed yield and HI, the opportunity to identify genetic loci that are important regulators in these complex traits is available. For example, a QTL analysis between L*ER* and Ob-0 could be carried out as these two ecotypes differ greatly in their seed yield.

Although a substantial decrease in seed yield per plant was found with increasing planting density, there was no effect on seed yield, total biomass and HI of the whole plant population. Previous experiments using field-grown crop plants and a wider range of densities had found alterations in these traits (Donald and Hamblin, 1976). Therefore, including more extreme population densities (such as 94, 188, and 3,000 plants/m^2) in our scaled-down *Arabidopsis* experiment could reveal more information. For instance, the effect of population density on seed size could be confirmed. Seed size in these experiments was determined by weighing the mass of 3 batches of 50 seeds and calculating the mean seed weight. This method does not take into account the variation of seed sizes produced within the population. It is possible that plants grown under higher population densities, and therefore higher stress, will produce a wider range of seed sizes than plants grown at lower population densities. For this reason, it would be interesting to determine

the variation in seed sizes produced. This could be carried out by passing the mature seeds through wire mesh of progressively smaller apertures.

A detailed analysis of seed development in restricted and unrestricted pollinations was performed in order to understand further the mechanism of increased seed size that occurs when seed number is restricted. The seed coat of seeds from restricted pollinations must be increased to accommodate a larger embryo and endosperm. However, only a small increase in the number of cells in the seed coat was detected, indicating that the seed coat cells undergo further cell elongation. The length of seed coat cells should therefore be determined in order to confirm that both cell division and cell expansion contribute to the increased size of the seed coat in restricted pollinations. Comparatively few modifications to seed development were found in the large seeds produced in restricted pollinations compared to the large seeds produced in interploidy crosses between a diploid mother plant and tetraploid father plant. To continue this study, a detailed analysis of the metabolic events that occur in restricted and unrestricted pollinations should be carried out as metabolic factors, such as sugars, have been shown to have a role in determining seed size (Weber *et al*., 1996, 1998; Focks and Benning, 1998; Hill *et al*., 2003). In addition, identification of the genes associated with the uplift in seed size in restricted pollinations by microarray analysis may provide potential tools for increasing seed size.

6.3.6 Summary

The experiments presented in this chapter clearly show that although not a crop plant, *Arabidopsis* can be extremely useful for studying seed yield and its components. Considerable variation in seed size, seed yield and HI has been shown to exist in *Arabidopsis*. As in many crop plants, large seed size was not associated with high yield or high HI. However, high seed number and reduced plant stature were revealed as important components of high yield and yield efficiency. In addition, the *ERECTA* gene has been identified as a potential target for improving HI.

Seed size in *Arabidopsis* was shown to be an extremely plastic trait in response to alterations in the post-flowering source-sink ratio, increasing by 35.0 – 66.2% in the ecotypes tested. Furthermore, a difference in seed weight between restricted (high source-sink ratio) and unrestricted (low source-sink ratio) pollinations was first observed remarkably early in seed development. However, reducing seed number by restricting pollination did not substantially alter integument or endosperm development in order to facilitate increased nutrient uptake.

7. DISCUSSION

7.1 Seed size modification: a potential strategy for yield improvement

The work described in this thesis aimed to determine whether a biotechnological approach to increasing seed size could produce a potentially useful increase in seed yield. In the past however, conventional breeding techniques that attempted the same strategy to raise seed yield in crop species failed due to yield component compensation. Seed yield remained unchanged as increases in seed size were accompanied by a reduction in seed number (Hartwig and Edwards, 1970; White and Izquierdo, 1991). The analyses of seed size and yield in *Arabidopsis* presented here (Chapter 6) show that yield component compensation also exists in this model species, with a strong negative correlation between seed number and size. Furthermore, it has been demonstrated that large seed size is not associated with high seed yield. Given this evidence, it may seem contradictory to modify seed size in order to improve seed yield. However, strategies were designed and tested that targeted both endosperm-led (Chapter 3) and integument-led (Chapter 4) seed growth, with several positive outcomes.

Endosperm-specific over-expression of the *ANT* transcription factor led to an increase in seed weight in both restricted and unrestricted pollinations. Fertility (seeds/per silique) in plants containing the *pPER::ANT* transgene was not affected (reduced) suggesting a potential increase in yield. Yield improvement can now be assessed objectively by comparing total seed yield/plant and HI in transgenic *pPER::ANT* and wild-type plants. It has also been shown that a biotechnological approach to modify the ovule integuments has the potential to increase seed size and seed yield. Restoration of floral opening in *arf2* mutants with the *pAP1::ARF2* transgene improved fertility whilst maintaining increased seed size. Additionally, targeted suppression of *ARF2* in the gynoecium using the *pSHP2::ARF2 RNAi* transgene led to an increase in seed size although seed set was reduced, likely as a result of an over-elongated gynoecium. Further yield trials in *B. napus* are necessary to determine whether yield enhancement is achievable in this economically-important species, where pollination by wind and insects should reduce the impact on fertility of a disparity between the stamen and gynoecium lengths.

As described above, modification of the *ANT* and *ARF2* genes showed great potential for increasing seed size and yield, however altering the expression of other single genes did not have the desired results. Introduction of the *pPER::PHE1* construct into *Arabidopsis* in order to up-regulate *PHE1* gene expression specifically within the endosperm, did not increase seed size compared to wild-type plants. Seeds developing on

pPER::AGL28 plants had a larger embryo sac area than wild-type however final seed size was not significantly altered in either restricted or unrestricted pollinations. Increased *SUGAR* transcripts were detected in siliques of plants containing the *pCZE::SUG*AR construct and a significant effect on seed development was observed. Surprisingly, over-expression of this sugar transporter in the chalazal endosperm led to a decrease in seed size in restricted pollinations. Seed size in unrestricted pollinations was not altered in *pCZE::SUGAR* plants.

Recent literature provides further evidence that the trade-off between seed number and seed size can be overcome to enable an improvement in seed yield. Constitutive over-expression of *GIBBERELLIC ACID-STIMULATED ARABIDOPSIS 4* (*GASA4*) is associated with significant increases in both mean seed weight and total seed yield/plant (Roxrud *et al*., 2007). The *Arabidopsis DA1* (Chinese for *'LARGE 1'*) gene contributes to final seed and organ size through restriction of the cell proliferation period (Li *et al*., 2008). *da1-1* mutants produced seeds that were 32% heavier than wild-type, had normal fertility and as a result the total seed yield was increased. In rice, loss-of-function of the *GW2* QTL region conferred a significant increase in seed size due to increased cell numbers and seed filling, ultimately enhancing total seed yield by 19.7% (Song *et al*., 2007).

Modifications to seed development that increase seed size can therefore, in some instances, improve seed yield. This evidence supports the concept that seed growth is, at least in part, controlled by the seed itself and therefore that the seeds themselves have an important role in the yield production process (Egli, 1998, 2006). In plants successfully engineered to produce large seeds and greater seed yield, the seeds must have been modified in such a way that increases sink strength and assimilate supply. The greater supply of assimilates to seeds can only have been achieved through a higher photosynthetic rate or altered partitioning (Egli, 1998). Measurement of the HI can provide clues as to which mechanism is responsible for the increased resource supply.

Alterations to the post-flowering source-sink ratio have been shown to cause dramatic changes in seed size in *Arabidopsis* (Chapter 6). With only minor changes in seed development, seeds increased in size by 35.0% to 66.2%, demonstrating a high level of plasticity in this trait. This suggests that under normal growth conditions (unrestricted pollinations) seeds have the capacity to grow much larger than is actually achieved. Reducing seed number by restricting pollination did not substantially alter seed development in order to facilitate resource uptake; instead the availability of more resources simply stretched the seed to full capacity. From this, it is difficult to ascertain whether seeds have a passive or active role in determining the extent of seed fill and final

seed size in response to source-sink alterations. Under conditions of high competition (unrestricted pollinations), the seeds may be attempting to acquire their full share of resources but are competing with equally aggressive siblings and so remain 'small'. When competition is reduced (restricted pollinations), seeds may be able to actively take on more resources, and as a result grow larger. The alternative could involve the mother plant force-feeding her offspring, with the seed playing a more passive role. Whichever mechanism is responsible, seed size and yield in *Arabidopsis* appears primarily source-limited.

7.2 Arabidopsis: a model for studying yield

Since *Arabidopsis* is not a crop species and has no direct economic importance, the question of its suitability as a model for studying seed yield and testing yield improvement strategies arises. In this thesis it has been shown that yield component relationships are strikingly similar in *Arabidopsis* and many crop species (Chapter 6): a negative trade-off between seed number and seed size is present in *Arabidopsis* and has been previously observed in sorghum (Kiniry, 1988) and maize (Kiniry *et al*., 1990; Borrás and Otegui, 2001); seed size is not associated with yield in *Arabidopsis* or many cereal crops (Peltonen-Sainio *et al*., 2007); and seed number is positively correlated with yield in *Arabidopsis*, as well as soybean, wheat (Egli, 1998), maize (Jong *et al*., 1982) and sunflower (Cantagallo *et al*., 1997). This suggests that a successful strategy for increasing yield in *Arabidopsis* may be transferable to crops.

As *Arabidopsis* is a dicot, the information gained from this model will no doubt be more relevant to dicotyledonous than to monocotyledonous crops. The divergence of these two clades is an ancient event, occurring between 140 and 235 million years ago (Yang *et al*., 1999; Chaw *et al*., 2004). *Arabidopsis* provides a particularly good model for the major oilseed crop *B. napus* (dicot), which is primarily grown for food and feed and has recently developed further interest as a source of biodiesel. The *Arabidopsis* and Brassica lineages diverged only 20 million years ago (Yang *et al*., 1999) and comparison of the nucleotide sequences showed that coding regions are highly conserved, having on average 85% similarity (Schmidt, 2002). The *B. napus* homologue of *AtARF2* has been identified (Schruff, 2006) and presents an opportunity to investigate whether this gene affects seed size and yield in this crop species.

Recently, the popularity of rice (*Oryza sativa*) as a model for monocot crops has risen, partly due to the complete sequencing of two widely-grown subspecies, indica and japonica (International Rice Genome Sequencing Project, 2005). Rice is unique as a model species since it is an economically important crop in its own right. A large disparity in gene number exists between *Arabidopsis* (25,000) and rice (40,000), although more than

80% of *Arabidopsis* genes have close homologues in rice (Bennetzen *et al*., 2004). However, the effects of many of these genes appear to be quite different in the two species. For example, constitutive over-expression of *DWARF4* (*DWF4*), which encodes an enzyme that catalyses the rate-limiting step in brassinosteroid biosynthesis, increases vegetative growth and seed yield in *Arabidopsis* (Choe *et al*., 2001). Contrastingly in rice, a similar approach decreased plant growth and seed production (Van Camp, 2005), whilst an *osdwf4-1* mutant improved yield due to leaves that were more erect, enhancing light capture for photosynthesis (Sakamoto *et al*., 2006). However, over-expression of *AtDWF4* in tobacco led to a similar phenotype as observed in *Arabidopsis* (Choe *et al*., 2001).

Due to the complexity of yield, it is extremely difficult to understand the underlying molecular and genetic basis of this trait. *Arabidopsis* remains the most efficient and low cost model and should continue to provide useful insight into the improvement of both monocot and dicot crops.

The current interest in growing crops for biofuels demands a more sophisticated approach to yield assessment that considers the total recoverable energy content. For example, homozygous *arf2-9* mutants produced a greater total mass (biological yield) than wild-type despite producing a dramatically lower seed yield, but a realistic evaluation of yield will require a comparison of total energy content. Such trials could be conducted in *Arabidopsis* provided that the various inputs to yield are recognised and adequately controlled.

7.3 Moving on: a combined strategy

No single *Arabidopsis* ecotype tested possessed all desirable yield-related characteristics (Chapter 6). Combining high seed weight, yield and HI into a single 'ideotype' is extremely difficult using conventional breeding. An 'ideotype' is an idealised plant type with a specific set of desirable characteristics to enhance photosynthesis, growth and seed production based on an understanding of plant growth, morphology and physiology (Donald, 1968). The ideotype approach to plant breeding was proposed as an alternative to 'defect elimination' and 'selection for yield'. Here, we show that *Arabidopsis* can be used effectively as a model for yield trials. Importantly, all the tools are available in *Arabidopsis* for building and testing a combined strategy through genetic modification.

It has recently been shown that seed yield/plant can be improved by the modification of single genes (Smidansky *et al*., 2002, 2003; Heyer *et al*., 2004; Ait-Ali *et al*., 2003). These findings indicate, along with other published data, that using single genes to increase seed size through modification of the endosperm (e.g. *ANT*) or integument (e.g. *ARF2*) remains a viable strategy to enhance seed yield (discussed in Section 7.1). *ERECTA*

has also been identified as a potential target to modify plant architecture and potentially improve HI. Since seed number was shown to have a strong positive correlation with seed yield, single genes associated with increased seed number/silique may provide a further target for crop improvement. A future EMS mutagenesis screen could aim to discover genes associated with this trait.

Achieving sufficient yield increases in the future may require the combination of several single gene modifications in one ideotype, targeting parameters such as seed size, seed number, seed composition, plant height, leaf number, leaf surface area and many more. Yield could be improved by the introduction of a particular set of useful gene expression cassettes, specific to each farmer-preferred variety according to local climate, consumer demands and agricultural practices. Ultimately, it must be remembered that yield is not simply the weight of grain per unit area but is more realistically the sale price of the harvested material minus the cost of production. The crop ideotype will require fewer inputs and give greater output (Donald, 1968).

8. REFERENCES

Adams, M.W. (1967) Basis of yield component compensation in crop plants with special reference to the field bean, *Phaseolus vulgaris*. *Crop Science*, **7:** 505-510

Adams, S., Vinkenoog, R., Spielman, M., Dickinson, H.G. and Scott, R.J. (2000) Parent-of-origin effects on seed development in *Arabidopsis thaliana* require DNA methylation. *Development,* **127**: 2493-2502

Adams, S. (2002) Genomic imprinting in the endosperm of *Arabidopsis thaliana*. *PhD Thesis*, University of Bath, UK

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**: 653-657

Ait-ali, T., Rands, C. and Harberd, N.P. (2003) Flexible control of plant architecture and yield via switchable expression of *Arabidopsis gai*. *Plant Biotechnology Journal,* **1**: 337-343

Alonso-Blanco, C., Blankestijn-De Vries, H., Hanhart, C.J. and Koornneef, M. (1999) Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. *Proceedings of the National Academy of Science USA,* **96**: 4710-4717

Arumuganathan, K. and Earle, E.D. (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter,* **9**: 211-215

Austin, R.B., Bingham, J., Blackwell, R.D., Evans, L.T., Ford, M.A., Morgan, C.L., and Taylor, M. (1980) Genetic improvements in winter wheat yields since 1900 and associated physiological changes. *Journal of Agricultural Science*, **94**: 675-689

Azpiroz, R., Wu, Y., Locascio, J.C. and Feldmann, K.A. (1998) An *Arabidopsis* brassinosteroid-dependent mutant is blocked in cell elongation. *The Plant Cell,* **10**: 219-230

Baker, S.C., Robinsonbeers, K., Villanueva, J.M., Gaiser, J.C. and Gasser, C.S. (1997) Interactions among genes regulating ovule development in *Arabidopsis thaliana*. *Genetics,* **145**: 1109-1124

Barker, G.C., Larson, T.R., Graham, I.A., Lynn, J.R. and King, G.J. (2007) Novel insights into seed fatty acid synthesis and modification pathways from genetic diversity and quantitative trait loci analysis of the *Brassica* C genome. *Plant Physiology,* **144**: 1827-1842

Barth, S., Busimi, A.K., Friedrich Utz, H. and Melchinger, A.E. (2003) Heterosis for biomass yield and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. *Heredity,* **91**: 36-42

Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L. and Rochat, C. (2002) An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiology and Biochemistry,* **40**: 151-160

Baud, S., Wuilleme, S., Lemoine, R., Kronenberger, J., Caboche, M., Lepiniec, L. and Rochat, C. (2005) The AtSUC5 sucrose transporter specifically expressed in the endosperm is involved in early seed development in *Arabidopsis*. *The Plant Journal,* **43**: 824-836

Baud, S., Dubreucq, B., Miquel, M., Rochat, C. and Lepiniec, L. (2008) Storage reserve accumulation in Arabidopsis: metabolic and developmental control of seed filling. *The Arabidopsis Book. American Society of Plant Biologists*. doi: 10.1199/tab.0113

Bazzaz, F.A., Carlson, R.W. and Harper, J.L. (1979) Contribution to reproductive effort by photosynthesis of flowers and fruits. *Nature,* **279**: 554-555

Beeckman, T., De Rycke, R., Viane, R. and Inze, D. (2000) Histological study of seed coat development in *Arabidopsis thaliana*. *Journal of Plant Research,* **113**: 139-148

Begna, S.H., Hamilton, R.I., Dwyer, L.M., Stewart, D.W. and Smith, D.L. (1997) Effects of population density on the yield and yield components of leafy reduced-stature maize in short-season areas. *Journal of Agronomy and Crop Science,* **178**: 103-110

Bennetzen, J.L., Coleman, C., Liu, R., Ma, J. and Ramakrishna, W. (2004) Consistent over-estimation of gene number in complex plant genomes. *Current Opinion in Plant Biology*, **7**: 732-736

Berleth, T. and Jürgens, G. (1993) The role of the *MONOPTEROS* gene in organising the basal body region of the *Arabidopsis* embryo. *Development,* **118**: 575-587

Borrás, L. and Otegui, M.E. (2001) Maize kernel weight response to post-flowering source-sink ratio. *Crop Science,* **49**: 1816-1822

Borrás, L., Westgate, M.E. and Otegui, M.E. (2003) Control of kernel weight and kernel water relations by post-flowering source-sink ratio in maize. *Annals of Botany,* **91**: 857-867

Borrás, L., Slafer, G.A. and Otegui, M.E. (2004) Seed dry weight response to source-sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Research,* **86**, 131-146

Braselton, J.P., Wilkinson, M.J. and Clulow, S.A. (1996) Feulgen staining of intact plant tissues for confocal microscopy. *Biotechnic and Histochemistry*, **71**: 84-87

Browse, J., Mccourt, P.J. and Somerville, C.R. (1986) Fatty-Acid Composition of Leaf Lipids Determined after Combined Digestion and Fatty-Acid Methyl-Ester Formation from Fresh Tissue. *Analytical Biochemistry,* **152**: 141-145

Cantagallo, J.E., Chimenti, C.A. and Hall, A.J. (1997) Number of seeds per unit area in sunflower correlates well with a photothermal quotient. *Crop Science*, **37**: 1780-1786

Cavell, A.C., Lydiate, D.J., Parkin, I.A.P., Dean, C. and Trick, M. (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. *Genome*, **41**: 62-69

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R. and Goodrich, J. (2004) Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. *Development,* **131**: 5263-5276

Chaw, S.M., Chang, C.C., Chen, H.L. and Li, W.H. (2004) Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution*, **58**: 424-441

Cheng, W.H., Taliercio, E.W. and Chourey, P.S. (1996) The *MINIATURE1* seed locus of maize encodes a cell wall invertase required for normal development of endosperm and maternal cells in the pedicel. *The Plant Cell,* **8**: 971-983

Choe, S., Fujioka, S., Noguchi, T., Takatsuto, S., Yoshida, S. and Feldmann, K.A. (2001) Overexpression of *DWARF4* in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in *Arabidopsis*. *The Plant Journal,* **26**: 573-582

Chuang, C.-F. and Meyerowitz, E.M. (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Science USA,* **97**: 4985-4990

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal,* **16**: 735-743

Cochrane, M.P. and Duffus, C.M. (1983) Endosperm cell number in cultivars of barley differing in grain weight. *Annals of Applied Biology*, **102**: 177-181

de Folter, S., Immink, R.G., Kieffer, M., Pařenicová, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B. and Angenent, G.C. (2005) Comprehensive interaction map of the *Arabidopsis* MADS Box transcription factors. *The Plant Cell,* **17**: 1424-1433

Debeaujon, I., Léon-Kloosterziel, K.M. and Koorneef, M. (2000) Influence of the testa on seed dormancy, germination and longevity in *Arabidopsis*. *Plant Physiology,* **122**: 403-413

Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M. and Lepiniec, L. (2003) Proanthocyanidin-accumulating cells in *Arabidopsis* testa: regulation of differentiation and role in seed development. *The Plant Cell,* **15**: 2514-2531

Department for Environment, Food and Rural Affairs (**DEFRA**), Arable crops grown in the UK, www.defra.gov.uk/farm/crops/index.htm#arable (accessed Sept. 2008)

Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005) The F-box protein TIR1 is an auxin receptor. *Nature,* **435**: 441-445

Dinneny, J.R., Weigel, D. and Yanofsky, M.F. (2005) A genetic framework for fruit patterning in *Arabidopsis thaliana*. *Development,* **132**: 4687-4696

Donald, C.M. (1962) In search of yield. *Journal of the Australian Institute of Agricultural Science,* **28**: 171-178

Donald, C.M. (1968) The breeding of crop ideotypes. *Euphytica*, **17**: 385-403

Donald, C.M. and Hamblin, J. (1976) Biological yield and harvest index of cereals as agronomic and plant breeding criteria. *Advances in Agronomy,* **28**: 361-405

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, **19**: 1349-1349

Egli, D.B. (1990) Seed water relations and the regulation of the duration of seed growth in soybean. *Journal of Experimental Botany*, **41**: 243-248

Egli, D.B. (1998) *Seed biology and the yield of grain crops*. CAB International, Wallingford, UK.

Egli, D.B. (2006) The role of seed in the determination of yield of grain crops. *Australian Journal of Agricultural Research*, **57**: 1237-1247

Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P. and Smyth, D.R. (1996) *AINTEGUMENTA*, an *APETALA2*-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *The Plant Cell,* **8**: 155-168

Ellis, C.M., Nagpal, P., Young, J.C., Hagen, G., Guilfoyle, T.J. and Reed, J.W. (2005) *AUXIN RESPONSE FACTOR1* and *AUXIN RESPONSE FACTOR2* regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development,* **132**: 4563-4574.

Eshed, Y., Baum, S.F. and Bowman, J.L. (1999) Distinct mechanisms promote polarity establishment in carpels of *Arabidopsis*. *Cell,* **99**: 199-209

Eshed, Y., Baum, S.F., Perea, J.V. and Bowman, J.L. (2001) Establishment of polarity in lateral organs of plants. *Current Biology,* **11**: 1251-1260Eshed *et al.*, 2001

Eshel, G. and Martin, P.A. (2006) Diet, energy, and global warming. *Earth Interactions,* **10**: 1-17

Evans, L.T. (1993) *Crop Evolution, Adaptation and Yield*. Cambridge University Press, Cambridge, UK.

Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A.R. and Galili, G. (2006) *Arabidopsis* seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiology,* **142**: 839-854

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X. and Zhang, Q. (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics,* **112**: 1164-1171

Ferrandiz, C., Pelaz, S. and Yanofsky, M.F. (1999) Control of carpel and fruit development in *Arabidopsis*. *Annual Review of Biochemistry,* **68**: 321-354

Flintham, J.E., Borner, A., Worland, A.J. and Gale, M.D. (1997) Optimising wheat grain yield: effects of Rht (gibberellin-insensitive) dwarfing genes. *Journal of Agricultural Science*, **128**: 11-25

Focks, N. and Benning, C. (1998) *wrinkled1*: a novel, low-seed-oil mutant of *Arabidopsis* with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiology*, **118**: 91-101

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N. and Snyder, P.K. (2005) Global consequences of land use. *Science,* **309**: 570-574

Food and Agricultural Organisation for the United Nations (**FAO**). Reducing poverty and hunger: the critical role of financing for food, agriculture and rural development. ftp://ftp.fao.org/docrep/fao/003/y6265e/y6265e.pdf (Data report from 2002, accessed March, 2009)

Food and Agricultural Organisation for the United Nations (**FAO**), FAOSTAT, faostat.fao.org/site/567/default.aspx#ancor (Data report from 2007, accessed September 2008)

Fridborg, I., Kuusk, S., Moritz, T. and Sundberg, E. (1999) The *Arabidopsis* dwarf mutant *shi* exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. *The Plant Cell,* **11**: 1019-1032

Gaiser, J.C., Robinson-Beers, K. and Gasser, C.S. (1995) The *Arabidopsis SUPERMAN* gene mediates asymmetric growth of the outer integument of ovules. *The Plant Cell,* **7**: 333-345

Gambín, B.L. and Borrás, L. (2007) Plasticity of sorghum kernel weight to increased assimilate availability. *Field Crops Research,* **100**: 272-284

Garcia, M.L., Saingery, V., Chambrier, P., Mayer, U., Jürgens, G. and Berger, F. (2003) *Arabidopsis haiku* mutants reveal new controls of seed size by endosperm. *Plant Physiology,* **131**: 1661-1670

Garcia, D., Fitz Gerald, J.N. and Berger, F. (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in *Arabidopsis*. *The Plant Cell,* **17**: 52-60

Gifford, M.L., Dean, S. and Ingram, G.C. (2003) The *Arabidopsis ACR4* gene plays a role in cell layer organisation during ovule integument and sepal margin development. *Development,* **130**: 4249-4258

Gonzalez, N., Beemster, G.T.S. and Inzé, D. (2008) David and Goliath: what can the tiny weed *Arabidopsis* teach us to improve biomass production in crops? *Current Opinion in Plant Biology*, **12**: 1-8

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L. and Henikoff, S. (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics,* **164**: 731-740

Groot, E.P. and Van Caeseele, A. (1993) The development of the aleurone layer in canola (*Brassica napus*). *Canadian Journal of Botany*, **71**: 1193-1201

Guilfoyle, T., Hagen, G., Ulmasov, T. and Murfett, J. (1998) How does auxin turn on genes? *Plant Physiology,* **118**: 341-347

Guilfoyle, T.J. and Hagen, G. (2007) Auxin response factors. *Current Opinion in Plant Biology,* **10**: 453-460

Haig, D. and Westoby, M. (1989) Parent-specific gene expression and the triploid endosperm. *The American Naturalist*, **134**: 147-155

Haig, D. and Westoby, M. (1991) Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Philosophical Transactions of the Royal Society of London*, **333**: 1-13

Hardtke, C.S. and Berleth, T. (1998) The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. *The EMBO Journal,* **17**: 1405-1411

Harper, J.L., Lovell, P.H. and Moore, K.G. (1970) The shapes and sizes of seeds. *Annual Review of Ecology and Systematics,* **1**: 327-356

Harper, R.M., Stowe-Evans, E.L., Luesse, D.R., Muto, H., Tatematsu, K., Watahiki, M.K., Yamamoto, K. and Liscum, E. (2000) The *NPH4* locus encodes the auxin response factor *ARF7*, a conditional regulator of differential growth in aerial *Arabidopsis* tissue. *The Plant Cell,* **12**: 757-770

Harrison, S.J., Mott, E.K., Parsley, K., Aspinall, S., Gray, J.C. and Cottage, A. (2006) A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. *Plant Methods*, **2**: 19-25

Hartwig, E.E. and Edwards, J. (1970) Effects of morphological characteristics upon seed yield in soybeans. *Agronomy Journal*, **62**: 64-65

Hauser, B.A., Villanueva, J.M. and Gasser, C.S. (1998) *Arabidopsis TSO1* regulates directional processes in cells during floral organogenesis. *Genetics,* **150**: 411-423

Hay, R.K.M. (1995) Harvest index - a review of its use in plant breeding and crop physiology. *Annals of Applied Biology,* **126**: 197-216

Heath, J.D., Weldon, R., Monnot, C., and Meinke, D.W. (1986) Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*. *Planta*, **169**: 304-312

Heyer, A.G., Raap, M., Schroeer, B., Marty, B. and Willmitzer, L. (2004) Cell wall invertase expression at the apical meristem alters floral, architectural, and reproductive traits in *Arabidopsis thaliana*. *The Plant Journal,* **39**: 161-169

Hill, L.M., Morley-Smith, E.R. and Rawsthorne, S. (2003) Metabolism of sugars in the endosperm of developing seeds of oilseed rape. *Plant Physiology,* **131**: 228-236

Hill, J. (2007) Environmental costs and benefits of transportation biofuel production from food- and lignocellulose-based energy crops. A review. *Agronomy for Sustainable Development,* **27**: 1-12

Huang, J., Pray, C. and Rozelle, S. (2002) Enhancing the crops to feed the poor. *Nature,* **418**: 678-684

Huang, X., Qian, Q., Liu, Z., Sun, H., He, S., Luo, D., Xia, G., Chu, C., Li, J. and Fu, X. (2009) Natural variation at the *DEP1* locus enhances grain yield in rice. *Nature Genetics,* **41**: 494-497

Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M. and Yamaguchi, J. (2001) *slender rice*, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. *The Plant Cell,* **13**: 999-1010

International Data Base (**IDB**), U.S. Census Bureau: World Population Information, http://www.census.gov/ipc/www/idb/worldpop.html (accessed Aug. 2008)

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature*, **436**: 793-800

James, C. (2005) Global status of commercialized biotech/GM crops: 2005. *ISAAA Briefs* No. 34. *International Service for the Acquisition of Agri-biotech Applications*, Ithaca, NY.

Jander, G., Baerson, S.R., Hudak, J.A., Gonzalez, K.A., Gruys, K.J. and Last, R.L. (2003) Ethylmethanesulfonate saturation mutagenesis in *Arabidopsis* to determine frequency of herbicide resistance. *Plant Physiology,* **131**: 139-146

Jofuku, K.D., Denboer, B.G.W., Vanmontagu, M. and Okamuro, J.K. (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. *The Plant Cell,* **6**: 1211-1225

Jofuku, K.D., Omidyar, P.K., Gee, Z. and Okamuro, J.K. (2005) Control of seed mass and seed yield by the floral homeotic gene *APETALA2*. *Proceedings of the National Academy of Science USA,* **102**: 3117-3122

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. *Science*, **290**: 344-347

Johnson, C.S., Kolevski, B. and Smyth, D.R. (2002) *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *The Plant Cell,* **14**: 1359-1375

Jong, S.K., Brewbaker, J.L. and Lee, C.H. (1982) Effects of solar radiation on the performance of maize in 41 successive monthly plantings in Hawaii. *Crop Science*, **22**: 13-18

Kang, I.H., Steffen, J.G., Portereiko, M.F., Lloyd, A. and Drews, G.N. (2008) The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. *The Plant Cell,* **20**: 635-647

Keith, K., Kraml, M., Dengler, N.G. and McCourt, P. (1994) *fusca3*: a heterochronic mutation affecting late embryo development in *Arabidopsis*. *The Plant Cell*, **6**: 589-600

Kepinski, S. and Leyser, O. (2005) The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature,* **435**: 446-451

Khush, G. (1999) Green Revolution: preparing for the 21st century. *Genome,* **42**, 646-655

King, S.P., Lunn, J.E. and Furbank, R.T. (1997) Carbohydrate content and enzyme metabolism in developing canola siliques. *Plant Physiology,* **114**: 153-160

Kiniry, J.R. (1988) Kernel weight increase in response to decreased kernel number in sorghum. *Agronomy Journal*, **80**: 221-226

Kiniry, J.R., Wood, C.A., Spanel, D.A. and Bockholt, A.J. (1990) Seed weight response to decreased seed number in maize. *Agronomy Journal*, **54**: 98-102

Klucher, K.M., Chow, H., Reiser, L. and Fischer, R.L. (1996) The *AINTEGUMENTA* gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. *The Plant Cell,* **8**: 137-153

Koelewijn, H.P. and Van Damme, J.M.M (2005) Effects of seed size, inbreeding and maternal sex on the offspring performance of gynodioecious *Plantago coronopus*. *Journal of Ecology*, 93: 373-383

Kohler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W. and Grossniklaus, U. (2003a) The Polycomb-group protein *MEDEA* regulates seed development by controlling expression of the MADS-box gene *PHERES1*. *Genes and Development*, **17**: 1540-1553

Kohler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U. and Gruissem, W. (2003b) *Arabidopsis* MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *The EMBO Journal,* **22**: 4804-4814

Kohler, C., Page, D.R., Gagliardini, V. and Grossniklaus, U. (2005) The *Arabidopsis thaliana MEDEA* Polycomb group protein controls expression of *PHERES1* by parental imprinting. *Nature Genetics,* **37**: 28-30

Koncz, C. and Schell, J. (1986) The promoter of Tl-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular & General Genetics,* **204**: 383-396

Koornneef, M. (1990) Mutations affecting the testa colour in *Arabidopsis*. *Arabidopsis Information Service*, **27**: 1-4

Krannitz, P.G., Aarssen, L.W. and Dow, J.M. (1991) The effect of genetically based differences in seed size on seedling survival in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany,* **78**: 446-450

Krizek, B.A. (1999) Ectopic expression *AINTEGUMENTA* in *Arabidopsis* plants results in increased growth of floral organs. *Developmental Genetics,* **25**: 224-236

Kudoh, H., Kachi, N., Kawano, S. and Ishiguri, Y. (2002) Intrinsic cost of delayed flowering in annual plants: negative correlation between flowering time and reproductive effort. *Plant Species Biology,* **17**: 101-107

Kusaba, M. (2004) RNA interference in crop plants. *Current Opinion in Biotechnology*, **15**: 139-143

Kyjovska, Z., Repkova, J. and Relichova, J. (2003) New embryo lethals in *Arabidopsis thaliana*: basic genetic and morphological study. *Genetica,* **119**: 317-325

Léon-Kloosterziel, K.M., Keijzer, C.J. and Koorneef, M. (1994) A seed shape mutant of *Arabidopsis* that is affected in integument development. *The Plant Cell,* **6**: 385-392

Li, H., Johnson, P., Stepanova, A., Alonso, J.M. and Ecker, J.R. (2004) Convergence of signaling pathways in the control of differential cell growth in *Arabidopsis*. *Developmental Cell,* **7**: 193-204

Li, Y.H., Beisson, F., Pollard, M. and Ohlrogge, J. (2006) Oil content of *Arabidopsis* seeds: the influence of seed anatomy, light and plant-to-plant variation. *Phytochemistry,* **67**: 904-915

Li, Y., Zheng, L., Corke, F., Smith, C. and Bevan, M.W. (2008) Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. *Genes and Development,* **22**: 1331-1336

Lopes, M.A. and Larkins, B.A. (1993) Endosperm origin, development, and function. *The Plant Cell*, **5**: 1383-1399

Lukowitz, W., Gillmor, C.S. and Scheible, W.R. (2000) Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiative working for you. *Plant Physiology,* **123**: 795-805

Luo, M., Dennis, E.S., Berger, F., Peacock, W.J. and Chaudhury, A. (2005) MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA,* **102**: 17531-17536

Luo, M., Luo, M.Z., Buzas, D., Finnegan, J., Helliwell, C., Dennis, E.S., Peacock, W.J. and Chaudhury, A. (2008) UBIQUITIN-SPECIFIC PROTEASE 26 is required for seed development and the repression of *PHERES1* in *Arabidopsis*. *Genetics,* **180**: 229-236

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U. and Kohler, C. (2006) Different Polycomb group complexes regulate common target genes in *Arabidopsis*. *EMBO Reports,* **7**: 947-952

Makarevich, G., Villar, C.B., Erilova, A. and Kohler, C. (2008) Mechanism of *PHERES1* imprinting in *Arabidopsis*. *Journal of Cell Science,* **121**: 906-912

Mallikarjuna Swamy, B.P. and Sarla, N. (2008) Yield-enhancing quantitative trait loci (QTLs) from wild species. *Biotechnology Advances,* **26**: 106-120

Malmberg, R.L. (2004) Plant mutagenesis and mutant screening. *Encyclopedia of Life Sciences, John Wiley & Sons, Ltd*

Mandel, M.A., Gustafsonbrown, C., Savidge, B. and Yanofsky, M.F. (1992) Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature,* **360**: 273-277

Masle, J., Gilmore, S.R. and Farquhar, G.D. (2005) The *ERECTA* gene regulates plant transpiration efficiency in *Arabidopsis*. *Nature,* **436**: 866-870

McAbee, J.M., Hill, T.A., Skinner, D.J., Izhaki, A., Hauser, B.A., Meister, R.J., Venugopala Reddy, G., Meyerowitz, E.M., Bowman, J.L. and Gasser, C.S. (2006) *ABERRANT TESTA SHAPE* encodes a KANADI family member, linking polarity determination to separation and growth of *Arabidopsis* ovule integuments. *The Plant Journal,* **46**: 522-531

Mizukami, Y. and Fischer, R.L. (2000) Plant organ size control: *AINTEGUMENTA* regulates growth and cell numbers during organogenesis. *Proceedings of the National Academy of Sciences of the USA,* **97**: 942-947

Mohamed-Yasseen, Y., Barringer, S.A., Splittstoesser, W.E. and Costanza, S. (1994) The role of seed coats in seed viability. *The Botanical Review*, **60**: 426-439

Morley-Smith, E.R., Pike, M.J., Findlay, K., Kockenberger, W., Hill, L.M., Smith, A.M. and Rawsthorne, S. (2008) The transport of sugars to developing embryos is not via the bulk endosperm in oilseed rape seeds. *Plant Physiology,* **147**: 2121-2130

Munier-Jolain, N.G. and Salon, C. (2005) Are the carbon costs of seed production related to the quantitative and qualitative performance? An appraisal for legumes and other crops. *Plant Cell and Environment,* **28**: 1388-1395

Murata, Y. (1969) Physiological responses to nitrogen in plants. In: Eastin, J.D., Haskins, F.A., Sullivan, C.Y. and Bavel, C.H.M.V. (eds) *Physiological aspects of crop yield*. American Society of Agronomy – Crop Science Society of America, Madison, Wisconsin.

Nagpal, P., Ellis, C.M., Weber, H., Ploense, S.E., Barkawi, L.S., Guilfoyle, T.J., Hagen, G., Alonso, J.M., Cohen, J.D., Farmer, E.E., Ecker, J.R. and Reed, J.W. (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development,* **132**: 4107-4118

Nelson, O.E. and Ohlrogge, A.J. (1957) Differential responses to population pressures by normal and dwarf lines of maize. *Science,* **125**: 1200

Nguyen, H., Brown, R.C. and Lemmon, B.E. (2000) The specialized chalazal endosperm in *Arabidopsis thaliana* and *Lepidium virginicum* (Brassicaceae). *Protoplasma,* **212**: 99-110

Ohto, M., Fischer, R.L., Goldberg, R.B., Nakamura, K. and Harada, J.J. (2005) Control of seed mass by *APETALA2*. *Proceedings of the National Academy of Sciences USA,* **102**: 3123-3128

Okushima, Y., Mitina, I., Quach, H.L. and Theologis, A. (2005) *AUXIN RESPONSE FACTOR 2* (*ARF2*): a pleiotropic developmental regulator. *The Plant Journal,* **43**: 29-46.

Orsi, C.H. and Tanksley, S.D. (2009) Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. *PLoS Genetics,* **5**: e1000347 doi:10.1371/journal.pgen.1000347

Olsen, O.-A. (2004) Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *The Plant Cell*, **16**: S214-S227

O'Neill, C.M., Gill, S., Hobbs, D., Morgan, C. and Bancroft, I. (2003) Natural variation for seed oil composition in *Arabidopsis thaliana*. *Phytochemistry,* **64**: 1077-1090

Page, D.R. and Grossniklaus, U. (2002) The art and design of genetic screens: *Arabidopsis thaliana*. *Nature Reviews Genetics,* **3**: 124-136

Pařenicová, L., De Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C. and Colombo, L. (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *The Plant Cell,* **15**: 1538-1551

Peltonen-Sainio, P., Kangas, A., Salo, Y. and Jauhiainen, L. (2007) Grain number dominates grain weight in temperate cereal yield determination: evidence based on 30 years of multi-location trials. *Field Crops Research,* **100**: 179-188

Peng, J.R., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D. and Harberd, N.P. (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature,* **400**: 256-261

Penning De Vries, F.W.T., Brunsting, A.H.M. and Van Laar, H.H. (1974) Products, requirements and efficiency of biosynthesis: a quantitative approach. *Journal of Theoretical Biology,* **45**: 339-377

Perry, M.W. and D'Antuono, M.F. (1989) Yield improvement and associated characteristics of some Australian spring wheat cultivars introduced between 1860 and 1982. *Australian Journal of Agricultural Research*, **40**: 457-472

Poneleit, C.G. and Egli, D.B. (1979) Kernel growth rate and duration in maize as affected by plant density and genotype. *Crop Science*, **19**: 385-388

Reiser, L. and Fischer, R.L. (1993) The ovule and the embryo sac. *The Plant Cell*, **5**: 1291-1301

Riou-Khamlichi, C., Menges, M., Healy, J.M.S. and Murray, J.A.H. (2000) Sugar control of the plant cell cycle: differential regulation of *Arabidopsis* D-type cyclin gene expression. *Molecular and Cellular Biology,* **20**: 4513-4521

Robinson, C.K. and Hill, S.A. (1999) Altered resource allocation during seed development in *Arabidopsis* caused by the *abi3* mutation. *Plant Cell and Environment,* **22**: 117-123

Robinson-Beers, K., Pruitt, R.E., and Gasser, C.S. (1992) Ovule development in wild-type *Arabidopsis* and two female-sterile mutants. *The Plant Cell*, **4**: 1237-1249

Roxrud, I., Lid, S.E., Fletcher, J.F., Schmidt, E.D.L. and Opsahl-Sorteberg, H.-G. (2007) GASA4, one of the 14-member *Arabidopsis* GASA family of small polypeptides, regulates flowering and seed development. *Plant and Cell Physiology,* **48**: 471-483

Sadras, V.O. (2007) Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Research,* **100**: 125-138

Sadras, V.O. and Egli, D.B. (2008) Seed size variation in grain crops: allometric relationships between rate and duration of seed growth. *Crop Science,* **48**: 408-416

Sakai, H., Medrano, L.J. and Meyerowitz, E.M. (1995) Role of *SUPERMAN* in maintaining *Arabidopsis* floral whorl boundaries. *Nature,* **378**: 199-203

Sakamoto, T. and Matsuoka, M. (2004) Generating high-yielding varieties by genetic manipulation of plant architecture. *Current Opinion in Biotechnology,* **15**: 144-147

Sakamoto, T., Morinaka, Y., Ohnishi, T., Sunohara, H., Fujioka, S., Ueguchi-Tanaka, M., Mizutani, M., Sakata, K., Takatsuto, S., Yoshida, S., Tanaka, H., Kitano, H. and Matsuoka, M. (2006) Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nature Biotechnology*, **24**: 105-109

Sasahara, T. (1984) Panicle properties and ripening, In: *Biology of rice*, Elsevier, Amsterdam pp 173-184

Savidge, B., Rounsley, S.D. and Yanofsky, M.F. (1995) Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. *The Plant Cell,* **7**: 721-733

Scott, R.J., Spielman, M., Bailey, J. and Dickinson, H.G. (1998) Parent-of-origin effects on seed development in *Arabidopsis thaliana*. *Development,* **125**: 3329-3341

Schmidt, R. (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Molecular Biology*, **48**: 21-37

Schmidt, R., Stransky, H., and Kock, W. (2007) The amino acid permease AAP8 is important for early seed development in *Arabidopsis thaliana*. *Planta*, **226**: 805-813

Schneitz, K., Hulskamp, M. and Pruitt, R.E. (1995) Wild-type ovule development in *Arabidopsis thaliana* - a light microscope study of cleared whole-mount tissue. *The Plant Journal,* **7**: 731-749

Schruff, M.C., Spielman, M., Tiwari, S., Adams, S., Fenby, N. and Scott, R.J. (2006) The *AUXIN RESPONSE FACTOR 2* gene of *Arabidopsis* links auxin signalling, cell division, and the size of seeds and other organs. *Development,* **133**: 251-261

Schruff, M.C. (2006) *megaintegumenta*, a seed size mutant of *Arabidopsis thaliana*. *PhD Thesis*, University of Bath, UK.

Sessions, R.A. and Zambryski, P.C. (1995) *Arabidopsis* gynoecium structure in the wild and in *ettin* mutants. *Development,* **121**: 1519-1532

Sessions, A., Nemhauser, J.L., Mccoll, A., Roe, J.L., Feldmann, K.A. and Zambryski, P.C. (1997) *ETTIN* patterns the *Arabidopsis* floral meristem and reproductive organs. *Development,* **124**: 4481-4491

Shapouri, S. and Rosen, S. (2007) Global diet composition: factors behind the changes and implications of the new trends. *Food Security Assessment, Economic Research Service*

Sheen, J., Zhou, L. and Jang, J.C. (1999) Sugars as signaling molecules. *Current Opinion in Plant Biology,* **2**: 410-418

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M. and Goodman, H.M. (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *The Plant Journal,* **8**: 659-671

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S. and Yano, M. (2008) Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genetics,* **40**: 1023-1028

Shpak, E.D., Lakeman, M.B. and Torii, K.U. (2003) Dominant-negative receptor uncovers redundancy in the *Arabidopsis* ERECTA leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *The Plant Cell,* **15**: 1095-1110

Shpak, E.D., Berthiaume, C.T., Hill, E.J. and Torii, K.U. (2004) Synergistic interaction of three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development,* **131**: 1491-1501

Sills, G.R. and Nienhuis, J. (1995) Maternal phenotypic effects due to soil nutrient levels and sink removal in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany*, **82**: 491-495

Sinclair, T.R. (1998) Historical changes in harvest index and crop nitrogen accumulation. *Crop Science,* **38**: 638-643

Smidansky, E.D., Clancy, M., Meyer, F.D., Lanning, S.P., Blake, N.K., Talbert, L.E. and Giroux, M.J. (2002) Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *Proceedings of the National Academy of Science USA,* **99**: 1724-1729

Smidansky, E.D., Martin, J.M., Hannah, L.C., Fischer, A.M. and Giroux, M.J. (2003) Seed yield and plant biomass increases in rice are conferred by deregulation of endosperm ADP-glucose pyrophosphorylase. *Planta,* **216**: 656-664

Smyth, D.R., Bowman, J.L. and Meyerowitz, E.M. (1990) Early flower development in *Arabidopsis*. *The Plant Cell,* **2**: 755-767

Somerville, C. and Koornneef, M. (2002) A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics,* **3**: 883-889

Song, X.J., Huang, W., Shi, M., Zhu, M.Z. and Lin, H.X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics,* **39**: 623-630

Stadler, R., Lauterbach, C. and Sauer, N. (2005) Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domains in *Arabidopsis* seeds and embryos. *Plant Physiology*, **139**: 701-712

Stowe-Evans, E.L., Harper, R.M., Motchoulski, A.V. and Liscum, E. (1998) *NPH4*, a conditional modulator of auxin-dependent differential growth responses in *Arabidopsis*. *Plant Physiology,* **118**: 1265-1275

Tajima, Y., Imamura, A., Kiba, T., Amano, Y., Yamashino, T. and Mizuno, T. (2004) Comparative studies on the type-B response regulators revealing their distinctive properties in the His-to-Asp phosphorelay signal transduction of *Arabidopsis thaliana. Plant and Cell Physiology* **45:** 28–39

Tan, X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M. and Zheng, N. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature,* **446**: 640-645

Tanaka, H., Watanabe, M., Watanabe, D., Tanaka, T., Machida, C. and Machida, Y. (2002) *ACR4*, a putative receptor kinase gene of *Arabidopsis thaliana*, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. *Plant and Cell Physiology,* **43**: 419-428

Tanska, M., Konopka, M. and Rotkiewicz, D. (2008) Relationships of rapeseed strength properties to seed size, colour and coat fibre composition. *Journal of the Science of Food and Agriculture,* **88**: 2186-2193

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**: 796-815

Thornley, J.H.M. (1980) Research strategy in the plant sciences. *Plant, Cell and Environment*, **3**: 233-236

Tiwari, S.B., Hagen, G. and Guilfoyle, T. (2003) The roles of auxin response factor domains in auxin-responsive transcription. *The Plant Cell,* **15**: 533-543

Tiwari, S., Spielman, M., Day, R.C. and Scott, R.J. (2006) Proliferative phase endosperm promoters from *Arabidopsis thaliana*. *Plant Biotechnology Journal,* **4**: 393-407

Tiwari, S., Spielman, M., Schulz, R., Oakey, R.J., Kelsey, G., Salazar, A., Zhang, K., Pennell, R. and Scott, R.J. (unpublished) Transcriptional profiles underlying parent-of-origin effects in seeds of *Arabidopsis thaliana*.

Tollenaar, M. and Wu, J. (1999) Yield improvement in temperate maize is attributable to greater stress tolerance. *Crop Science*, **39**: 1597-1604

Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F. and Komeda, Y. (1996) The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell*, **8**: 735-746

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1997) ARF1, a transcription factor that binds to auxin response elements. *Science,* **276**: 1865-1868

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1999) Dimerization and DNA binding of auxin response factors. *The Plant Journal,* **19**: 309-319

Ungerer, M.C., Halldorsdottir, S.S., Modliszewski, J.L., Mackay, T.F. and Purugganan, M.D. (2002) Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*. *Genetics,* **160**: 1133-1151

Ungru, A., Nowack, M.K., Reymond, M., Shirzadi, R., Kumar, M., Biewers, S., Grini, P.E. and Schnittger, A. (2008) Natural variation in the degree of autonomous endosperm formation reveals independence and constraints of embryo growth during seed development in *Arabidopsis thaliana*. *Genetics,* **179**: 829-841

Van Camp, W. (2005) Yield enhancement genes: seeds for growth. *Current Opinion in Biotechnology,* **16**: 147-153

Van Doorn, W.G. and Van Meeteren, U. (2003) Flower opening and closure: a review. *Journal of Experimental Botany,* **54**: 1801-1812

Van Zanten, M., Snoek, L.B., Proveniers, M.C. and Peeters, A.J. (2009) The many functions of ERECTA. *Trends in Plant Science,* **14**: 214-218

Vanneste, S. and Friml, J. (2009) Auxin: a trigger for change in plant development. *Cell,* **136**: 1005-1016

Vaughn, J.G. and Whitehouse, J.M. (1971) Seed structure and the taxonomy of the Cruciferae. *Botanical Journal of the Linnean Society*, **64**: 383-409

Venable, D.L. (1992) Size-number trade-offs and the variation of seed size with plant resource status. *The American Naturalist,* **140**: 287-304

Vert, G., Walcher, C.L., Chory, J. and Nemhauser, J.L. (2008) Integration of auxin and brassinosteroid pathways by *AUXIN RESPONSE FACTOR 2*. *Proceedings of the National Academy of Science USA,* **105**: 9829-9834

Vertregt, N. and Penning De Vries, F.W.T. (1987) A rapid method for determining the efficiency of biosynthesis of plant biomass. *Journal of Theoretical Biology,* **128**: 109-119

Vilhar, B., Kladnik, A., Blejec, A., Chourey, P.S. and Dermastia, M. (2002) Cytometrical evidence that the loss of seed weight in the *miniature1* seed mutant of maize is associated with reduced mitotic activity in the developing endosperm. *Plant Physiology*, **129**: 23-30

Villanueva, J.M., Broadhves, T.J., Hauser, B.A., Meister, R.J., Schneitz, K. and Gasser, C.S. (1999) INNER NO OUTER regulates abaxial–adaxial patterning in *Arabidopsis* ovules. *Genes & Development,* **13**: 3160-3169

Watanabe, M., Tanaka, H., Watanabe, D., Machida, C. and Machida, Y. (2004) The ACR4 receptor-like kinase is required for surface formation of epidermis-related tissues in *Arabidopsis thaliana*. *The Plant Journal,* **39**: 298-308

Weber, H., Borisjuk, L. and Wobus, U. (1996) Controlling seed development and seed size in *Vicia faba*: a role for seed coat-associated invertases and carbohydrate state. *The Plant Journal,* **10**: 823-834

Weber, H., Heim, U., Golombek, S., Borisjuk, L. and Wobus, U. (1998) Assimilate uptake and the regulation of seed development. *Seed Science Research,* **8**: 331-345

Weijers, D., Benkova, E., Jager, K.E., Schlereth, A., Hamann, T., Kientz, M., Wilmoth, J.C., Reed, J.W. and Jurgens, G. (2005) Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *The EMBO Journal,* **24**: 1874-1885

West, M. and Harada, J.J. (1993) Embryogenesis in higher plants: an overview. *The Plant Cell,* **5**: 1361-1369

Westgate, M.E. and Grant, D.L. (1989a) Water deficits and reproduction in maize: response of the reproductive tissue to water deficits at anthesis and mid-grain fill. *Plant Physiology,* **91**: 862-867

Westgate, M.E. and Grant, D.T. (1989b) Effect of water deficits on seed development in soybean: I. Tissue water status. *Plant Physiology,* **91**: 975-979

Westoby, M. and Rice, B. (1982) Evolution of the seed plants and inclusive fitness of plant tissues. *Evolution,* **36**: 713-724

Westoby, M., Leishman, M. and Lord, J. (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* **351**: 1309-1317

White, J.W. and Izquierdo, J. (1991) Physiology of yield potential and stress tolerance. In: vanSchoonhover, A. and Voysest, O. (eds) *Common beans: research for crop improvement.* CAB International, Wallingford, UK.

Windsor, J.B., Symonds, V.V., Mendenhall, J. and Lloyd, A.M. (2000) *Arabidopsis* seed coat development: morphological differentiation of the outer integument. *The Plant Journal,* **22**: 483-493

Wobus, U. and Weber, H. (1999) Sugars as signal molecules in plant seed development. *Biological Chemistry,* **380**: 937-944

Yang, Y.-W., Lai, K.-N., Tai, P.Y. and Li, W.H. (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *Journal of Molecular Evolution*, **48**: 597-604

Yeung, E.C. and Meinke, D.W. (1993) Embryogenesis in angiosperms: development of the suspensor. *The Plant Cell*, **5**: 1371-1381

Yoo, S.K., Lee, J.S. and Ahn, J.H. (2006) Overexpression of *AGAMOUS-LIKE 28* (*AGL28*) promotes flowering by upregulating expression of floral promoters within the autonomous pathway. *Biochemical and Biophysical Research Communications,* **348**: 929-936

Zhang, J.H. and Lechowicz, M.J. (1994) Correlation between time of flowering and phenotypic plasticity in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany,* **81**: 1336-1342

Zhou, Y., Setz, N., Niemietz, C., Qu, H., Offler, C.E., Tyerman, S.D. and Patrick, J.W. (2007) Aquaporins and unloading of phloem-imported water in coats of developing bean seeds. *Plant, Cell and Environment,* **30**: 1566-1577